Annals of Pure and Applied Logic 35 (1987) 261–302
North-Holland

ON THE SCHEME OF INDUCTION FOR BOUNDED ARITHMETIC FORMULAS

A.J. WILKIE and J.B. PARIS
Department of Mathematics, The University, Manchester M13 9PL, U.K.

Communicated by D. Van Dalen
Received 20 December 1985

1. Introduction

1.1. Let $L = \{\mathbf{0}, ', +, \cdot, \leqslant\}$ denote the usual first-order language for arithmetic. Recall that an L-formula is called Δ_0, or bounded, if all its quantifiers occur in context $\forall x\,(x \leqslant t(\boldsymbol{y}) \rightarrow \cdots)$ or $\exists x\,(x \leqslant t(\boldsymbol{y}) \wedge \cdots)$ (abbreviated to $\forall x \leqslant t(\boldsymbol{y}) \cdots$, $\exists x \leqslant t(\boldsymbol{y}) \cdots$ respectively), where t is a term of L. In this paper we investigate certain extensions of the system $\mathrm{I}\Delta_0$ (bounded induction) which is axiomatized in L by the scheme

$$\forall \boldsymbol{x}, z\,((\phi(\boldsymbol{x}, 0) \wedge \forall y \leqslant z\,(\phi(\boldsymbol{x}, y) \rightarrow \phi(\boldsymbol{x}, y'))) \rightarrow \forall y \leqslant z\,\phi(\boldsymbol{x}, y)) \quad (\phi \in \Delta_0)$$

together with sufficient elementary arithmetic properties of the basic symbols, for example the following will do:

$$(\mathbf{0} \leqslant \mathbf{0} \wedge \neg \mathbf{0}' \leqslant \mathbf{0}),$$
$$\forall x\,(x + \mathbf{0} = x \wedge x \cdot \mathbf{0} = \mathbf{0} \wedge x \cdot \mathbf{0}' = x),$$
$$\forall x\,\forall y\,(x' = y' \rightarrow x = y),$$
$$\forall x\,\forall y\,(x \leqslant y' \leftrightarrow (x \leqslant y \vee x = y')),$$
$$\forall x\,\forall y\,x + y' = (x + y)',$$
$$\forall x\,\forall y\,x \cdot y' = (x \cdot y) + x.$$

It has been known for some time (see [6]) that if $\phi(\boldsymbol{x}, y) \in \Delta_0$ and $\mathrm{I}\Delta_0 \vdash \forall \boldsymbol{x}\,\exists y\,\phi(\boldsymbol{x}, y)$, then for some term t of L we have $\mathrm{I}\Delta_0 \vdash \forall \boldsymbol{x}\,\exists y \leqslant t(\boldsymbol{x})\,\phi(\boldsymbol{x}, y)$. One consequence of this fact is that informal number-theoretic or combinatorial arguments which make use, even implicitly, of a function of non-polynomial growth cannot be translated directly into formal proofs using only the axioms of $\mathrm{I}\Delta_0$. Now if the conclusions of such arguments make no mention of the fast-growing functions used in their proof (e.g., if they are informal renderings of Π_1^0 sentences of L), then the question as to whether (or which) such statements are provable at all in $\mathrm{I}\Delta_0$ seems to us particularly interesting, especially in view of

the large number of natural examples. In order to make this question mor
precise (and to give some of the examples) we introduce some more notation.

In [4] it is shown that there is a Δ_0 formula, which we shall denote by $x^y = z$ which can be proved in $\mathrm{I}\Delta_0$ to have all the usual basic algebraic properties of th graph of the exponential function with the exception (because of the resul mentioned above) of $\forall x\, \forall y\, \exists z\, (x^y = z)$. We denote this last sentence of L b exp.

The system $\mathrm{I}\Delta_0 + \exp$ seems to be strong enough to carry out most of th proofs encountered in elementary number theory and combinatorics and it als has many natural equivalent systems, e.g., elementary function arithmetic (ope induction + defining equations for the Kalmar elementary functions). We ar therefore interested in to what extent, and how, the axiom exp can be eliminate from proofs in $\mathrm{I}\Delta_0 + \exp$ of Π_1^0 sentences of L. It will turn out that for a larg subclass of the Π_1^0 sentences, the use of exponentiation can be replaced by that c certain natural consistency statements and so we shall also be investigating suc statements for their own interest.

We now present two illustrations of the above comments, both of which giv rise to unsolved problems in this area and as such have already appeared in th literature (see [10], [11], [9], [12]).

1.2. The Euclid proof of the infinity of primes can easily be formalized i $\mathrm{I}\Delta_0 + \exp$, but clearly requires exp. Can this theorem be proved in $\mathrm{I}\Delta_0$ alone?

1.3. In [4] it is shown that Matijasevič's theorem on the equality of Σ_1^0-definabl and existentially definable sets is provable in $\mathrm{I}\Delta_0 + \exp$. Can exp be significantl weakened here? Any result in this direction would have considerable cons quences for low level complexity theory. (For further discussion on th connections between Δ_0 sets and complexity theory, which has partly motivate the investigations in this paper, see [7], [8], [9] and [11].)

1.4. The main results of this paper are stated in the survey article [9] (we sha restate them in Section 3), where much of the required notation can also b found. However, we have tried to keep this paper reasonably self-containe although we do require on behalf of the reader a knowledge of some bas theorems and constructions that are within the scope of $\mathrm{I}\Delta_0$ and certain of i extensions. Good references for this are [2] and [12].

2. Preliminaries

2.1. Let M be any model of $\mathrm{I}\Delta_0$. As usual we identify M with its domain an suppose that the natural numbers, ω, form an initial segment of M. For $n \in \omega$, w denote by $\mathbf{n}$ the term $\mathbf{0}\underbrace{'\cdots'}_{n}$. If α is a formula (respectively term, closed term) (

L, α^M denotes the set (respectively function, element) it defines in M. Thus $\mathbf{n}^M = n$. If a formula, $\phi(\boldsymbol{x}, y)$, of L is known to define a function (total or partial) in M, or in any theory under consideration, we shall usually use the more suggestive notation $\phi(\boldsymbol{x}) = y$ (as we did for "$x^y = z$" in Section 1) for $\phi(\boldsymbol{x}, y)$.

We shall often want to regard elements of M as ($M-$) finite sequences over some (actually) finite alphabet. It is usual to take $\{1, 2\}$ as this alphabet and use dyadic coding but we feel that it is more natural, especially for the encoding of logical syntax in M, to use a larger alphabet. So let us fix a large natural number B ($B = 30$ will certainly suffice) to be used throughout this paper as the number base. We shall use B-adic coding. Thus every non-zero $\alpha \in M$ may be 'written' uniquely as

$$\alpha = \sum_{i=0}^{t} \alpha_i \cdot B^i \qquad (t \in M,\ 1 \leqslant \alpha_i \leqslant B \text{ for } i = 0, \ldots, t)$$

and hence identified with the 'word' $\alpha_0\alpha_1 \cdots \alpha_t$ from the alphabet $\{1, 2, \ldots, B\}$. We also identify 0 with the empty word. The length of α, denoted $|\alpha|$, is $t+1$, and $|0| = 0$. The important point here is that the relations $|x| = y$ and $x_z = y$ ("y is the zth member of the word x") are defined by Δ_0 formulas, namely

$$(y = 0 \wedge x = 0) \vee \exists z \leqslant \mathbf{B} \cdot x (z = \mathbf{B}^y \wedge z \geqslant x + 1),$$

and

$$\mathbf{1} \leqslant y \leqslant \mathbf{B} \wedge z' \leqslant |x| \wedge \exists u \leqslant x\, \exists v \leqslant x\, (v = \mathbf{B}^z \wedge |u| \leqslant z \wedge \exists w \leqslant x\, (w \cdot v \cdot \mathbf{B} + u + y \cdot v = x)),$$

respectively.

We shall also denote by $xy = z$ ("the concatenation of x and y is z") the Δ_0 formula

$$\exists u \leqslant z\, \exists v \leqslant x\, (v = |x| \wedge u = \mathbf{B}^v \wedge z = x + u \cdot y),$$

so that if $\alpha, \beta \in M$, $\alpha = \alpha_0 \cdots \alpha_t$, $\beta = \beta_0 \cdots \beta_s$, then $\alpha\beta = \alpha_0 \cdots \alpha_t\beta_0 \cdots \beta_s$. It is easy to establish in $\mathrm{I}\Delta_0$ all the usual properties of these relations (using properties of the formula $x^y = z$ from [4]). We mention in particular

2.2. If α, β, $\gamma \in M$, then

$$M \vDash \exists!\, z\, (z = \alpha\beta) \wedge |\alpha\beta| = |\alpha| + |\beta| \wedge \alpha(\beta\gamma) = (\alpha\beta)\gamma \wedge \alpha\beta \leqslant \mathbf{B}^2 \cdot \alpha \cdot \beta.$$

2.3. If $\phi(x, y)$ is a Δ_0 formula (possibly involving parameters from M), $\alpha \in M$ and $M \vDash \forall x \leqslant \alpha\, \exists!\, y\, (1 \leqslant y \leqslant B \wedge \phi(x, y)) \wedge \exists z\, (z = B^\alpha)$, then there is a unique $\beta \in M$ such that

$$M \vDash \forall x \leqslant \alpha\, \forall y\, (\phi(x, y) \leftrightarrow y = \beta_x) \wedge |\beta| = \alpha + 1.$$

In view of 2.2 we shall often regard concatenation as a term when writing down formulas and also omit parentheses when concatenating several terms. We shall, therefore, never use the common convention of omitting the multiplication symbol '$\cdot$'. Thus, $\phi(xyz)$ is an abbreviation for the formula

$$\psi(x, y, z) \stackrel{\text{def}}{=} \exists u \leqslant \mathbf{B}^2 \cdot x \cdot y \, \exists v \leqslant \mathbf{B}^2 \cdot u \cdot z \, (u = xy \wedge v = uz \wedge \phi(v)).$$

Notice that if $\phi(v)$ is Δ_0, then so is ψ.

Since we are working to a base greater than ten, we must resolve an annoying ambiguity, which we do as illustrated in the following example. 14 denotes the number $1 + (4 \cdot B)$, i.e., the word consisting of the digit 1 followed by the digit 4. The number fourteen is denoted $\overline{14}$. Similarly, $\overline{\mathbf{14}}$ denotes the term $\mathbf{0}\underbrace{{}'' \cdots {}'}_{\text{fourteen times}}$, whereas $x = \mathbf{14}$ is the formula expressing "x is the result of concatenating $\mathbf{1}$ and $\mathbf{4}$".

2.4 For $n \in \omega$, the $n+1$-place function e_n is defined as follows:

$$e_0(x_1) = x_1, \qquad e_{n+1}(x_1, \ldots, x_{n+2}) = x_1^{e_n(x_2, \ldots, x_{n+2})}.$$

Thus

$$e_n(x_1, \ldots, x_{n+1}) = x_1^{x_2^{\cdot^{\cdot^{x_{n+1}}}}}.$$

We also set $\omega_n(x) = e_n(x, |x|, \|x\|, \ldots, |x|^{(n)})$ where $|x|^{(n)}$ denotes the result of applying the length function, $|\cdot|$, n times to x.

Clearly the relations $e_n(x_1, \ldots, x_{n+1}) = y$ and $\omega_n(x) = y$ for $n \in \omega$ can be expressed by Δ_0 formulas, but in $\mathrm{I}\Delta_0$ one can only show that they define partial functions. We denote the L-sentence $\forall x \, \exists y \, \omega_n(x) = y$, expressing the fact that ω_n is total, by Ω_n.

For much of this paper we shall work with the theory $\mathrm{I}\Delta_0 + \Omega_1$ rather than $\mathrm{I}\Delta_0$. This is because it allows more flexible constructions on words than does $\mathrm{I}\Delta_0$ alone. Indeed, the set of *lengths* of words in any model of $\mathrm{I}\Delta_0 + \Omega_1$ is closed under multiplication, whereas this set will in general only be closed under addition in a model of $\mathrm{I}\Delta_0$. (Thus, for example, $\mathrm{I}\Delta_0 + \Omega_1$ is the weakest theory in which it is sensible to discuss polynomial time computations, and hence the ?P = NP? question. For if $\alpha \in M \models \mathrm{I}\Delta_0 + \Omega_1$, then any polynomial time computation with input (the word) α will also be coded in M. In $\mathrm{I}\Delta_0$ alone it only seems possible to formalise linear time computations.) Perhaps we should also remark that in most respects $\mathrm{I}\Delta_0 + \Omega_1$ is much more akin to $\mathrm{I}\Delta_0$ than it is to $\mathrm{I}\Delta_0 + \exp$, so we feel the motivation mentioned in Section 1 for studying $\mathrm{I}\Delta_0$ holds good for $\mathrm{I}\Delta_0 + \Omega_1$.

2.5. We denote by $x \subseteq_{\mathrm{p}} y$ ("x is a part of y") the Δ_0 formula $\exists u, v \leqslant y \, (y = uxv)$. Consider the following conditions on a set S of formulas.

(1) $(\tau_1\tau_2 = \tau_3) \in S$, $(\tau_1\tau_2 \neq \tau_3) \in S$ for any terms τ_1, τ_2, τ_3, of L.

(2) $\phi \in S$ implies $\exists x \subseteq_{\mathrm{p}} y \, \phi \in S$ and $\forall x \subseteq_{\mathrm{p}} y \, \phi \in S$, where these formulas are abbreviations for $\exists x \, (x \subseteq_{\mathrm{p}} y \wedge \phi)$ and $\forall x \, (x \subseteq_{\mathrm{p}} y \rightarrow \phi)$ respectively.

(3) $\phi \in S$ and $\psi \in S$ imply $(\phi \wedge \psi) \in S$ and $(\phi \vee \psi) \in S$.

(4) $\phi \in S$ implies $\exists x\, (|x| \leqslant |y| \wedge \phi) \in S$. This formula is abbreviated to $\exists\, |x| \leqslant |y|\, \phi$.

(5) $\phi \in S$ implies $\exists x\, (|x| \leqslant t(|y|) \wedge \phi) \in S$, for any term $t(v)$ of L. This formula is abbreviated to $\exists\, |x| \leqslant t(|y|)\, \phi$.

(6) $\phi \in S$ implies $\neg\phi \in S$.

The smallest sets S of formulas satisfying, respectively, (1)–(3), (1)–(4), (1)–(5), (1)–(4) and (6), (1)–(6) are called SR (strictly rudimentary), R^+ (positive rudimentary), R_1^+ (extended positive rudimentary), RUD (rudimentary), RUD_1 (extended rudimentary).

It is easy to see that the formulas $\tau_1 = \tau_2$, $\tau_1 \neq \tau_2$, $\tau_1 \subseteq_{\mathrm{p}} x$, $\tau_1 \nsubseteq_{\mathrm{p}} x$ (i.e., $\neg\tau_1 \subseteq_{\mathrm{p}} x$) are all equivalent in $\mathrm{I}\Delta_0$ to SR formulas. Further, every RUD formula is equivalent in $\mathrm{I}\Delta_0$ to a Δ_0 formula (and, as is immediate, conversely) so we may use induction for RUD formulas in models of $\mathrm{I}\Delta_0$. It is also easy to see that the scheme of induction for RUD_1 formulas holds in any model of $\mathrm{I}\Delta_0 + \Omega_1$ (cf. the remarks in 2.4; see also [3], [8], [9] for detailed discussions on these classes of formulas).

In Section 4 we shall see that all syntactic notions can be formalised using R_1^+ formulas and it will be crucial that these notions are preserved from a model of $\mathrm{I}\Delta_0$ to certain of its extensions. Unfortunately, we do not know if R_1^+ formulas are preserved to arbitrary extensions of models of $\mathrm{I}\Delta_0$ (this problem is clearly related to 1.3), so we make the following

2.6. Definition. Suppose M_1, $M_2 \models \mathrm{I}\Delta_0$. Then we write $M_1 \subseteq_+ M_2$ if whenever $\phi(x_1, \ldots, x_n)$ is R_1^+ and $a_1, \ldots, a_n \in M_1$ and $M_1 \models \phi(a_1, \ldots, a_n)$, then

$$a_1, \ldots, a_n \in M_2 \quad \text{and} \quad M_2 \models \phi(a_1, \ldots, a_n).$$

Notice that $M_1 \subseteq_+ M_2$ implies $M_1 \subseteq M_2$ because atomic formulas of L are R_1^+.

2.7. Remark. When working in, say, Peano arithmetic (or even $\mathrm{I}\Delta_0 + \exp$) one usually only takes care to check that syntactic notions are defined by Σ_1^0 formulas. One reason for this is that if $\phi(x)$ is a Σ_1^0 formula, $a \in M$ (M any model under consideration) and $M \models \phi(a)$, then M will contain a proof of (a suitable formalization of) $\phi(a)$. This essentially follows from the fact that M contains a proof of the 'formula'

$$\forall x \left(x \leqslant a \rightarrow \bigvee_{i \leqslant a} x = i \right).$$

Now if M is not closed under exponentiation then, for suitable a, this formula will be too long to be coded in M. However, consider the formula

$$\forall x \left(x \subseteq_{\mathrm{p}} a \rightarrow \bigvee_{i \subseteq_{\mathrm{p}} a} x = i \right).$$

Its length is of order $|a|^2$, and so it will be coded, and in fact have a proof, in M provided $M \vDash \mathrm{I}\Delta_0 + \Omega_1$. This in turn will imply that M contains a proof of $\phi(a)$ for any R_1^+ formula ϕ such that $M \vDash \phi(a)$. Thus it is the R_1^+ formulas that, in the theories we are considering, will play the role usually enjoyed by the Σ_1^0 formulas.

There is also a comparison from the point of view of computability that can be made here. While the Σ_1^0 formulas define exactly the recursively enumerable sets in the standard model, the R_1^+ formulas define exactly the non-deterministic polynomial time computable sets (NP).

3. Statement of results

Our first theorem states that the theories $\mathrm{I}\Delta_0 + \Omega_n$, $n = 0, 1, 2, \ldots,$ are equiconsistent despite the fact that they clearly increase in strength as n increases.

3.1. Theorem. *For any $n \in \omega$,*

$$\mathrm{I}\Delta_0 + \Omega_1 + \mathrm{Con}(\mathrm{I}\Delta_0) \vdash \mathrm{Con}(\mathrm{I}\Delta_0 + \Omega_n).$$

(Con(X) *is the formalisation, to be made precise later, of "the set of sentences X is consistent in the predicate calculus".*)

We also have Gödel's second incompleteness theorem for $\mathrm{I}\Delta_0$:

3.2. Theorem. $\mathrm{I}\Delta_0 \nvdash \mathrm{Con}(\mathrm{I}\Delta_0)$. *In fact, for any $n \in \omega$,* $\mathrm{I}\Delta_0 + \Omega_n \nvdash \mathrm{Con}(\mathrm{I}\Delta_0)$.

In order to study the axiom exp we introduce a model-theoretic construction:

3.3. Theorem. *Suppose that for all $k \in \omega$, $M \vDash \mathrm{I}\Delta_0 + \Omega_1 + \mathrm{Con}(\mathrm{I}\Delta_0, k)$, M countable, where* Con(X, k) *is the formalisation of the statement "there is no proof of a contradiction from the set of formulas X which involves only substitution instances of formulas of length $\leqslant k$". Then there is a model K of* $\mathrm{I}\Delta_0 + \exp$ *such that $M \subseteq_+ K$.*

From this we deduce our theorem on the elimination of exp from proofs of certain Π_1^0 sentences.

3.4. Theorem. *Suppose $\phi(\boldsymbol{x}) \in \mathrm{R}_1^+$. Then* $\mathrm{I}\Delta_0 + \exp \vdash \forall \boldsymbol{x} \neg\phi(\boldsymbol{x})$ *if and only if for some $k \in \omega$,*

$$\mathrm{I}\Delta_0 + \Omega_1 + \mathrm{Con}(\mathrm{I}\Delta_0, k) \vdash \forall \boldsymbol{x} \neg\phi(\boldsymbol{x}).$$

Using 3.3 we can also improve considerably on 3.2:

3.5. Theorem. $\mathrm{I}\Delta_0+\exp\nvdash\mathrm{Con}(\mathrm{I}\Delta_0)$. *Indeed,* $\mathrm{I}\Delta_0+\exp\nvdash\mathrm{Con}(\mathrm{Q})$, *where* Q *is Robinson's arithmetic. Also,*

$$\mathrm{I}\Delta_0+\exp+\mathrm{Con}(\mathrm{I}\Delta_0)\nvdash\mathrm{Con}(\mathrm{I}\Delta_0+\exp).$$

We shall also show that $\mathrm{I}\Delta_0+\exp$ is not a Π_1^0 conservative extension of $\mathrm{I}\Delta_0+\Omega_1$. Indeed, we have

3.6. Theorem. *For each* $n\in\omega$, *there is an* R_1^+ *formula,* $\phi(x)$, *such that* $\mathrm{I}\Delta_0+\exp\vdash\forall x\neg\phi(x)$, *but* $\mathrm{I}\Delta_0+\Omega_n\nvdash\forall x\neg\phi(x)$.

4. Arithmetization of syntax

4.1. As already mentioned our aim here is to construct R_1^+ formulas expressing such notions as "x is a term", "x is a formula", "x is a proof" etc. For simplicity we suppose that the language to be coded contains the basic logical symbols $\neg$, $\rightarrow$ and $\forall$ only ($\wedge$, $\vee$, $\leftrightarrow$, $\exists$ being introduced as abbreviations in the usual way) and that the variables are $v_1, v_2, \ldots, v_n, \ldots$. We Gödel number the basic symbols using the alphabet $\{3, 4, \ldots, B\}$ (1 and 2 will be used as markers) as follows:

$$\begin{array}{ccccccccccccc} (&) & \neg & \rightarrow & \forall & v & {}_1 & ' & + & \cdot & \leqslant & \mathbf{0} & = \\ 3 & 4 & 5 & 6 & 7 & 8 & 9 & \overline{10} & \overline{11} & \overline{12} & \overline{13} & \overline{14} & \overline{15}, \end{array}$$

and our intention is that the Gödel number of a formula (or term) will be its natural B-adic code. Thus, for example, the Gödel number of

$$\forall v_3((v_3'=\mathbf{0})\rightarrow((v_1+v_1)=\mathbf{0}'))$$

is obtained by first writing v_3 as v_{111} and then computing the number corresponding to the resulting word in the basic symbols, which in this case is $7+8\cdot B+9\cdot B^2+9\cdot B^3+9\cdot B^4+\cdots+\overline{15}\cdot B^{20}+\overline{14}\cdot B^{21}+\overline{10}\cdot B^{22}+4\cdot B^{23}+4\cdot B^{24}$.

That the sets of Gödel numbers of terms, formulas and, indeed, all syntactic notions can be defined by R_1^+ formulas follows from the fact that they are defined (at least, in usual presentations of the predicate calculus) by so-called 'recursion on notation' and R_1^+ is (provably in $\mathrm{I}\Delta_0+\Omega_1$) closed under this operation. We shall illustrate why this is by giving a detailed account of the definition of "x is a term", and then leave the reader to fill in the details for the other definitions. (Our results are not optimal in the sense that we could find R^+ (rather than R_1^+) definitions of these syntactic notions. However, the definitions we shall exhibit are very natural and are sufficient for our purposes.)

We first define some auxiliary R_1^+ formulas.

4.2. Definitions

(1) $\mathrm{Var}(x) \overset{\mathrm{def}}{\Leftrightarrow} \exists y_1\subseteq_{\mathrm{p}}x\,(x=\mathbf{8}y_1\wedge\forall y_2\subseteq_{\mathrm{p}}y_1\,(\mathbf{9}\subseteq_{\mathrm{p}}y_2))$.

(2) $\mathrm{Z}(x) \overset{\mathrm{def}}{\Leftrightarrow} x = \overline{\mathbf{14}}$.

(3) $\mathrm{Bterm}(x) \overset{\mathrm{def}}{\Leftrightarrow} (\mathrm{Var}(x) \vee \mathrm{Z}(x))$.

(4) $\mathrm{Suc}(x) = y \overset{\mathrm{def}}{\Leftrightarrow} y = x\overline{\mathbf{10}}$.

(5) $\mathrm{Sum}(x_1, x_2) = y \overset{\mathrm{def}}{\Leftrightarrow} y = \mathbf{3}x_1\overline{\mathbf{11}}x_2\mathbf{4}$.

(6) $\mathrm{Prod}(x_1, x_2) = y \overset{\mathrm{def}}{\Leftrightarrow} y = \mathbf{3}x_1\overline{\mathbf{12}}x_2\mathbf{4}$.

To define Term(x) ("x is a term") the idea is to assert the existence of a element s of the form $2z_12z_2 \cdots 2z_t2$ which enumerates the subterms of x. That is s has the property that for each $l = 1, \dots, t$, either (i) Bterm(z_l) or (ii $\exists j < l\, \mathrm{Suc}(z_j) = z_l$ or (iii) $\exists j, k < l\, \mathrm{Sum}(z_j, z_k) = z_l$ or (iv) $\exists j, k < l\, \mathrm{Prod}(z_j, z_k) = z_l$; and $z_t = x$. To avoid ambiguity we must also insist that each z_l neither contain a 2 in its B-adic expansion nor is 0. Thus we arrive at the following

4.3. Definitions

(1) $\mathrm{Termseq}(x) \overset{\mathrm{def}}{\Leftrightarrow} \mathbf{22} \not\subseteq_{\mathrm{p}} x \wedge \exists y \subseteq_{\mathrm{p}} x\, (y \neq \mathbf{0} \wedge \mathbf{2}y\mathbf{2} = x) \wedge \forall y \subseteq_{\mathrm{p}} x$
$(\mathbf{2} \subseteq_{\mathrm{p}} y \vee \mathbf{2}y\mathbf{2} \not\subseteq_{\mathrm{p}} x \vee \mathrm{Bterm}(y) \vee \exists x_1, y_1 \subseteq_{\mathrm{p}} x\, (\mathbf{2} \not\subseteq_{\mathrm{p}} x_1$
$\wedge \mathbf{2}x_1\mathbf{2}y_1y \subseteq_{\mathrm{p}} x \wedge \mathrm{Suc}(x_1) = y) \vee \exists x_1, x_2, y_1, y_2, z_1, z_2 \subseteq_{\mathrm{p}} x$
$(\mathbf{2} \not\subseteq_{\mathrm{p}} x_1 \wedge \mathbf{2} \not\subseteq_{\mathrm{p}} x_2 \wedge y_1 = \mathbf{2}z_1 \wedge y_1 = z_2\mathbf{2} \wedge \mathbf{2}x_1y_1x_2\mathbf{2}y_2y \subseteq_{\mathrm{p}} x$
$\wedge (\mathrm{Sum}(x_1, x_2) = y \vee \mathrm{Prod}(x_1, x_2) = y)))$.

(2) $\mathrm{Term}(x) \overset{\mathrm{def}}{\Leftrightarrow} \mathbf{2} \not\subseteq_{\mathrm{p}} x \wedge \exists y\, \mathrm{Termseq}(y\mathbf{2}x\mathbf{2})$.

Notice that Termseq(x) is R_1^+ (in fact SR) but Term(x) is not, at least as i stands, because of the unbounded existential quantifier appearing in th definition. However, it turns out that the subterm sequence of a given term ca be chosen to be not too much longer than the term itself. More precisely we hav

4.4. Lemma

$$\mathrm{I}\Delta_0 + \Omega_1 \vdash \forall x\, (\mathrm{Term}(x) \leftrightarrow (\mathbf{2} \not\subseteq_{\mathrm{p}} x \wedge \exists\, |y| \leq |x|^2\, \mathrm{Termseq}(y\mathbf{2}x\mathbf{2}))).$$

Thus we may assume Term(x) *is an* R_1^+ *formula.*

Proof. Consider the formula

$$\forall\, |x| \leq |t|\, ((\mathbf{2} \not\subseteq_{\mathrm{p}} x \wedge \exists\, |y| \leq |t|\, \mathrm{Termseq}(y\mathbf{2}x\mathbf{2})) \rightarrow (\mathbf{2} \not\subseteq_{\mathrm{p}} x \wedge \exists\, |u| \leq |x|^2\, \mathrm{Termseq}(u\mathbf{2}x\mathbf{2})))$$

which we denote by $\psi(t)$.

$\psi(t)$ is certainly RUD_1 and so, by the remarks in 2.5, we may use induction on t to establish $\forall t\, \psi(t)$ which clearly proves the lemma.

Work in $\mathrm{I}\Delta_0 + \Omega_1$. $\psi(\mathbf{0})$ is trivial, so suppose $t > 0$, $\psi(z)$ holds for all $z < t$ and that $|x| \leq |t|$, $|y| \leq |t|$, $\mathbf{2} \not\subseteq_{\mathrm{p}} x$ and $\mathrm{Termseq}(y\mathbf{2}x\mathbf{2})$. By 4.3(1) we have either $\mathrm{Bterm}(x)$, in which case $u = \mathbf{0}$ certainly satisfies $|u| \leq |x|^2 \wedge \mathrm{Termseq}(u\mathbf{2}x\mathbf{2})$ or for some x_1, $x_2 \subseteq_{\mathrm{p}} x$ (with $\mathbf{2} \not\subseteq_{\mathrm{p}} x_1, x_2$) we have either (a) $\mathrm{Suc}(x_1) = x$, or (b) $\mathrm{Sum}(x_1, x_2) = x$, or (c) $\mathrm{Prod}(x_1, x_2) = x$, and further, definition 4.3(1) clearly implies that for some $|y_1| < |y|$, $|y_2| < |y|$, we have $\mathrm{Termseq}(y_1\mathbf{2}x_1\mathbf{2})$ and $\mathrm{Termseq}(y_2\mathbf{2}x_2\mathbf{2})$. Since from 4.2(4), (5), (6) we also have $|x_1| < |x|$ and $|x_2| < |x|$, we may use the inductive hypothesis to deduce the existence of u_1, u_2 with $|u_i| < |x_i|^2$ and $\mathrm{Termseq}(u_i\mathbf{2}x_i\mathbf{2})$ $(i = 1, 2)$. Set

$$u = \begin{cases} u_1\mathbf{2}x_1 & \text{in case (a),} \\ u_1\mathbf{2}x_1u_2\mathbf{2}x_2 & \text{in case (b) or (c).} \end{cases}$$

Then clearly $\mathrm{Termseq}(u\mathbf{2}x\mathbf{2})$ and, in case (a),

$$|u| = \mathbf{1} + |u_1| + |x_1| \leq \mathbf{1} + |x_1|^2 + |x_1| \leq (|x_1| + \mathbf{1})^2 = |x|^2$$

(by 4.2(4)); and, in case (b) or (c),

$$|u| = \mathbf{2} + \sum_{i=1}^{2} (|u_i| + |x_i|) \leq \sum_{i=1}^{2} (|x_i| + \mathbf{1})^2 \leq (|x_1| + |x_2| + \mathbf{2})^2 \leq |x|^2$$

(by 4.2(5) or (6)), which completes the proof. □

4.5. Suppose $M \models \mathrm{I}\Delta_0 + \Omega_1$. Notice that if $m \in \omega$, then $M \models \mathrm{Term}(m)$ if and only if m is the Gödel number of a standard term, τ say, computed as described in 4.1, and we shall write $\ulcorner \tau \urcorner = m$ in this situation. Notice also that Term^M contains Bterm^M and is closed under Suc^M, Sum^M and Prod^M. Moreover, it is easy to show (by induction in M) that if A is a RUD_1 definable (even with parameters) subset of M containing Bterm^M and closed under Suc^M, Sum^M and Prod^M, then $\mathrm{Term}^M \subseteq A$. Thus it is reasonable to call $\mathrm{Term}(x)$ "the smallest RUD_1 formula containing Bterm and closed under Suc, Sum and Prod", and we shall indeed use analogous expressions in making subsequent definitions.

4.6. Definitions

(1) $\mathrm{Atform}(x) \overset{\text{def}}{\Leftrightarrow} \exists y_1, y_2 \subseteq_{\mathrm{p}} x\, (\mathrm{Term}(y_1) \wedge \mathrm{Term}(y_2)$
$\wedge\, (x = \mathbf{3}y_1\overline{\mathbf{13}}\,y_2\mathbf{4} \vee x = \mathbf{3}y_1\overline{\mathbf{15}}\,y_2\mathbf{4}))$.

(2) $\mathrm{Neg}(x) = y \overset{\text{def}}{\Leftrightarrow} y = \mathbf{5}x$.

(3) $\mathrm{Imp}(x_1, x_2) = y \overset{\text{def}}{\Leftrightarrow} y = \mathbf{3}x_1\mathbf{6}x_2\mathbf{4}$.

(4) $\mathrm{UQ}(x_1, x_2) = y \overset{\text{def}}{\Leftrightarrow} \exists y_1 \subseteq_{\mathrm{p}} x_1\, (\mathrm{Var}(y_1) \wedge x_1 = y_1\overline{\mathbf{15}}\,\overline{\mathbf{14}} \wedge y = \mathbf{7}y_1x_2)$.

(5) $\mathrm{Form}(x) \overset{\mathrm{def}}{=}$ "the smallest RUD_1 formula containing Atform and closed under Neg, Imp and UQ".

Thus Form(x) ("x is a formula") is obtained by first defining Formseq(x) by analogy with 4.3(1) and then existentially quantifying as in 4.3(2). A lemma corresponding to 4.4 is then easily established, showing that Form(x) is R_1^+. Remarks similar to those in 4.5 also apply, in particular we write $\ulcorner\phi\urcorner = m$ for "m is the Gödel number of the (standard) formula ϕ".

4.7. Definitions

(1) $\mathrm{BUQ}(x) = y \overset{\mathrm{def}}{\Leftrightarrow} \exists y_1, y_2, y_3 \subseteq_{\mathrm{p}} x\ (x = \mathbf{33}y_1\overline{\mathbf{13}}y_2\mathbf{46}y_3\mathbf{4}$

$\wedge \mathrm{Var}(y_1) \wedge \mathrm{Term}(y_2) \wedge \mathrm{Form}(y_3) \wedge y = \mathbf{7}y_1x).$

(2) $\Delta_0\text{-Form}(x) \overset{\mathrm{def}}{=}$ "the smallest RUD_1 formula containing Atform and closed under Neg, Imp and BUQ".

(3) $\mathrm{Clterm}(x) \overset{\mathrm{def}}{=}$ "the smallest RUD_1 formula containing Z (see 4.2(2)) and closed under Suc, Sum and Prod".

(4) For $i = 1, \ldots, B$, $\mathrm{Dig}_i(y) \overset{\mathrm{def}}{\Leftrightarrow} y = \overline{\mathbf{14}}\underbrace{\mathbf{10}\cdots\mathbf{10}}_{i\ \text{times}}$.

(5) For $i = 1, \ldots, B$, $\mathrm{In}_i(x) = y \overset{\mathrm{def}}{\Leftrightarrow} \exists y_1, y_2 \subseteq_{\mathrm{p}} y\ (\mathrm{Dig}_i(y_i)$

$\wedge \mathrm{Dig}_B(y_2) \wedge y = \mathbf{3}y_1\overline{\mathbf{11}}\,\mathbf{3}y_2\,\overline{\mathbf{12}}x\,\mathbf{44}).$

(6) $\mathrm{Canclterm}(x) \overset{\mathrm{def}}{=}$ "(the smallest RUD_1 formula containing $\mathrm{Dig}_1, \mathrm{Dig}_2, \ldots, \mathrm{Dig}_B$ and closed under $\mathrm{In}_1, \mathrm{In}_2, \ldots, \mathrm{In}_B) \vee (x = \overline{\mathbf{14}})$".

Again, the formula Δ_0-Form(x) ("x is a Δ_0 formula"), Clterm(x) ("x is a closed term") and Cancltterm(x) ("x is a canonical closed term") are formally defined by analogy with the definition of Term(x), and they are all R_1^+. Also one can easily establish in $\mathrm{I}\Delta_0 + \Omega_1$ the sentences

$$\forall x\ ((\mathrm{Cancltterm}(x) \to \mathrm{Clterm}(x)) \wedge (\mathrm{Clterm}(x) \to \mathrm{Term}(x)))$$

and

$$\forall x\ (\mathrm{Clterm}(x) \leftrightarrow (\mathrm{Term}(x) \wedge \forall y \subseteq_{\mathrm{p}} x\ (y \neq \mathbf{8}))).$$

(The canonical closed terms are those of the form $(\mathbf{a}_0 + (\mathbf{B} \cdot (\mathbf{a}_1 + \mathbf{B} \cdot (\mathbf{a}_2 + \cdots + (\mathbf{B} \cdot \mathbf{a}_t)) \cdots)$, where $1 \leq a_i \leq B$ for $i = 0, \ldots, t$, or $\mathbf{0}$.)

We now come to some more complicated definitions, namely those connected with the substitution of terms for variables. These are usually made by the 'principle of induction for formulas (or terms)', and we illustrate how this can be directly translated into our setting in one case, leaving the reader to check the details of the others. We first require two lemmas.

4.8. Lemma (Unique readability of terms and formulas)

(1) $\mathrm{I}\Delta_0 + \Omega_1 \vdash \forall x, y, z\ ((\mathrm{Term}(x) \wedge yz = x \wedge \mathrm{Term}(z)) \rightarrow y = \mathbf{0})$.

(2) $\mathrm{I}\Delta_0 + \Omega_1 \vdash \forall x, y, z\ ((\mathrm{Form}(x) \wedge yz = x \wedge \mathrm{Form}(y)) \rightarrow z = \mathbf{0})$.

Proof. For (1), suppose $M \vDash \mathrm{I}\Delta_0 + \Omega_1$ and work in M. We show, by induction on c, that if $ab = c$, $\mathrm{Term}(b)$ and $\mathrm{Term}(c)$ then $a = 0$.

This is trivial for $c = 0$. Suppose the result holds for $c < d$ and $ab = d$, $\mathrm{Term}(b)$ and $\mathrm{Term}(d)$. If either $\mathrm{Bterm}(b)$ or $\mathrm{Bterm}(d)$ the required result follows easily by inspection. Otherwise (by definition of Termseq) each of b and d are of one of the forms $e\,\overline{10}$, $3e\,\overline{11}\,f4$, $3e\,\overline{12}\,f4$, where $\mathrm{Term}(e)$, $\mathrm{Term}(f)$, $e < d$ and $f < d$. Now if one of b, d is of form $e\,\overline{10}$, then the other must be (since they must end in the same digit), and we have, say, $ae_1\overline{10} = e_2\overline{10}$ so $ae_1 = e_2$ and, by the inductive assumption, $a = 0$. Otherwise, we have, say, $a3e_1pf_14 = 3e_2qf_24$ where $p, q \in \{\overline{11}, \overline{12}\}$. Hence $a3e_1pf_1 = 3e_2qf_2$. It follows that either $a^*f_1 = f_2$ or $a^*f_2 = f_1$ (for some a^*), so by the inductive assumption we have $a^* = 0$ and $f_1 = f_2$. Hence $p = q$ and $a3e_1 = 3e_2$. A similar argument shows $e_1 = e_2$, so $a = 0$ as required.

The proof of (2) is similar. □

Using 4.8 and the ideas used in its proof it is now easy to establish the following lemma. The details are left to the reader.

4.9. Lemma. *Provably in* $\mathrm{I}\Delta_0 + \Omega_1$, Suc, Sum *and* Prof, *when restricted to* $\mathrm{Term}(x)$, *define one-one functions with pairwise disjoint images. Further each of these images is disjoint from* $\mathrm{Bterm}(x)$. *A similar result holds for* Neg, Imp, UQ *and* $\mathrm{Form}(x)$, $\mathrm{Atform}(X)$.

4.10. Definition

$$\mathrm{Presubt}(z_1, z_2, x) = y \overset{\text{def}}{\Leftrightarrow} \mathrm{Term}(z_1) \wedge \mathrm{Var}(z_2) \wedge \mathrm{Bterm}(x)$$
$$\wedge ((x = z_2 \wedge y = z_1) \vee (x \neq z_2 \wedge y = x)).$$

Thus $\mathrm{Presubt}(z_1, z_2, x) = y$ defines "y is the result of substituting the term z_1 for the variable z_2 in the basic term x", and we wish to extend this to "arbitrary terms x". In other words we require an R_1^+ formula, $\mathrm{Subt}(z_1, z_2, x) = y$, with $\mathrm{Subt}(z_1, z_2, x)$ defined whenever $\mathrm{Term}(z_1)$, $\mathrm{Var}(z_2)$ and $\mathrm{Term}(x)$ and satisfying (provably in $\mathrm{I}\Delta_0 + \Omega_1$) the equations

(1) $\mathrm{Bterm}(x) \rightarrow (\mathrm{Subt}(z_1, z_2, x) = \mathrm{Presubt}(z_1, z_2, x))$,

(2) $\mathrm{Term}(x) \rightarrow (\mathrm{Subt}(z_1, z_2, \mathrm{Suc}(x)) = \mathrm{Suc}(\mathrm{Subt}(z_1, z_2, x))$,

(3)
(4) $$(\mathrm{Term}(x_1) \wedge \mathrm{Term}(x_2)) \rightarrow \left(\mathrm{Subt}(z_1, z_2, \begin{smallmatrix}\mathrm{Sum}\\ \mathrm{Prod}\end{smallmatrix}(x_1, x_2)\right)$$
$$= \begin{smallmatrix}\mathrm{Sum}\\ \mathrm{Prod}\end{smallmatrix}(\mathrm{Subt}(z_1, z_2, x_1), \mathrm{Subt}(z_1, z_2, x_2)).$$

We use the same idea as before in that we first define Subtseq(z_1, z_2, s) by: s is of the form $\mathbf{2}u_1\mathbf{1}w_1\mathbf{2}\cdots\mathbf{2}u_t\mathbf{1}w_t\mathbf{2}$, satisfying, for $l=1,\ldots,t$

$$\begin{aligned}
&\text{either } \mathrm{Bterm}(u_l) \wedge w_l = \mathrm{Presubt}(z_1, z_2, u_l))\\
&\text{or } (\mathrm{Term}(u_l) \wedge \exists j<l\,(u_l = \mathrm{Suc}(u_j) \wedge w_l = \mathrm{Suc}(w_j)))\\
&\text{or } (\mathrm{Term}(u_l) \wedge \exists j,k<l\,(u_l = \mathrm{Sum}(u_j,u_k) \wedge w_l = \mathrm{Sum}(w_j,w_k)))\\
&\text{or } (\mathrm{Term}(u_l) \wedge \exists j,k<l\,(u_l = \mathrm{Prod}(u_j,u_k) \wedge w_l = \mathrm{Prod}(w_j,w_k))),
\end{aligned}$$

and then define Subt$(z_1, z_2, x) = y$ by

$$\mathrm{Term}(z_1) \wedge \mathrm{Var}(z_2) \wedge \mathrm{Term}(x) \wedge \exists s\ \mathrm{Subtseq}(z_1, z_2, s\mathbf{2}x\mathbf{1}y\mathbf{2}).$$

It is routine to write out a formula R_1^+ definition of Subtseq(z_1, z_2, s) and we leave this to the reader. However, in order to check that Subt is equivalent in $\mathrm{I}\Delta_0 + \Omega_1$ to an R_1^+ formula we must find a bound for the "$\exists s$" appearing in the definition and this will come from the fact that substitution does not increase length by too much. We must also show that Subt defines a partial function. Both of these goals are achieved in the following

4.11. Lemma

(1) $\mathrm{I}\Delta_0 + \Omega_1 \vdash \forall z_1,z_2,x\ ((\mathrm{Term}(z_1) \wedge \mathrm{Var}(z_2) \wedge \mathrm{Term}(x))$
$\rightarrow \exists\ |y| \leqslant (|z_1|\cdot|x|)\ \exists s \leqslant (|z_1|\cdot|x|^2)\ (\mathrm{Term}(y) \wedge \mathrm{Subtseq}(z_1, z_2, s\mathbf{2}x\mathbf{1}y\mathbf{2})))$.

(2) $\mathrm{I}\Delta_0 + \Omega_1 \vdash \forall z_1, z_2, x\ ((\mathrm{Term}(z_1) \vee \mathrm{Var}(z_2) \wedge \mathrm{Term}(x))$
$\rightarrow \forall y_1,y_2,s_1,s_2((\mathrm{Subtseq}(z_1, z_2, s_1\mathbf{2}x\mathbf{1}y_1\mathbf{2})$
$\wedge\ \mathrm{Subtseq}(z_1, z_2, s_2\mathbf{2}x\mathbf{1}y_2\mathbf{2})) \rightarrow y_1 = y_2))$.

Proof. (1) is proved in a similar way to 4.4, the only non-trivial step in the induction being to show that if $|y_1| \leqslant |z_1|\cdot|x_1|$, $|y_2| \leqslant |z_1|\cdot|x_2|$, $|s_1| \leqslant |z_1|\cdot|x_1|^2$, $|s_2| \leqslant |z_1|\cdot|x_2|^2$, $x = \mathrm{Suc}(x_1)$ or $\mathrm{Sum}(x_1, x_2)$ or $\mathrm{Prod}(x_1, x_2)$ and $y = \mathrm{Suc}(y_1)$ or $\mathrm{Sum}(y_1, y_2)$ or $\mathrm{Prod}(y_1, y_2)$ (respectively), then, with $s = s_1\mathbf{2}x_1\mathbf{1}y_1$ (in the first case) or with $s = s_1\mathbf{2}x_1\mathbf{1}y_1s_2\mathbf{2}x_2\mathbf{1}y_2$ (in the second and third cases), we have $|y| \leqslant |z_1|\cdot|x|$ and $|s| \leqslant |z_1|\cdot|x|^2$. This is an easy computation; for example if $x = \mathrm{Sum}(x_1, x_2)$ and $y = \mathrm{Sum}(y_1, y_2)$, then

$$\begin{aligned}
|y| &= |\mathbf{3}y_1\overline{\mathbf{11}}\,y_2\mathbf{4}| = \mathbf{3} + |y_1| + |y_2| \leqslant \mathbf{3} + |z_1|\cdot|x_1| + |z_1|\cdot|x_2|\\
&\leqslant |z_1|\cdot(\mathbf{3} + |x_1| + |x_2|) = |z_1|\cdot|x|,
\end{aligned}$$

and

$$\begin{aligned}
|s| &= \mathbf{4} + \sum_{j=1}^{2} (|s_j| + |y_j| + |x_j|)\\
&\leqslant \mathbf{4} + |z_1|\cdot\sum_{j=1}^{2} (|x_j|^2 + \mathbf{2}\cdot|x_j|)\\
&\leqslant \mathbf{2} + |z_1|\cdot\sum_{j=1}^{2} (|x_j| + 1)^2 \leqslant \mathbf{2} + |z_1|\cdot\Big(\sum_{j=1}^{2} |x_j| + 1\Big)^2\\
&= \mathbf{2} + |z_1|\cdot(|x| - \mathbf{1})^2 \leqslant |z_1|\cdot|x|^2 \qquad (\text{since } |x| \geqslant \mathbf{3}).
\end{aligned}$$

The computations for the other possibilities for x, y are similar.

(2) follows by induction on $t = \max(s_1\mathbf{2}x\mathbf{1}y_1\mathbf{2}, s_2\mathbf{2}x\mathbf{1}y_2\mathbf{2})$ with a crucial use of 4.9. For that lemma implies that exactly one of Bterm(x), $x = \mathrm{Suc}(x_1)$, $x = \mathrm{Sum}(x_1, x_2)$, $x = \mathrm{Prod}(x_1, x_2)$ holds, where x_1 and x_2 are uniquely determined by x and hence they must both occur as parts of s_1 and of s_2. This now allows us to reduce t and apply the inductive hypothesis. □

4.12. If $M \models \mathrm{I}\Delta_0 + \Omega_1$ then, by 4.11, Subt^M is an R_1^+-definable partial function with domain $\mathrm{Term}^M \times \mathrm{Var}^M \times \mathrm{Term}^M$ and range included in Term^M satisfying 4.10(1)–(4). It is easy to show by induction in M, that if f is any RUD_1-definable function with these properties then $f = \mathrm{Subt}^M$. Hence we call Subt "the unique RUD_1 definable function satisfying 4.10(1)–(4)" and use a similar expression in subsequent definitions. Note that we also have $\mathrm{Subt}^M(\ulcorner\sigma\urcorner, \ulcorner x\urcorner, \ulcorner\tau\urcorner) = \ulcorner\tau(x/\sigma)\urcorner$ for standard terms τ, σ and variables x.

4.13. Definition. Subf(z_1, z_2, x) ("the result of substituting the term z_1 for each free occurrence of the variable z_2 in the formula x") is the unique RUD_1 definable function, with the obvious domain, satisfying the following:

(1) $\mathrm{Atform}(x) \to (\mathrm{Subf}(z_1, z_2, x) = y \leftrightarrow (\mathrm{Term}(z_1) \wedge \mathrm{Var}(z_2)$
$$\wedge\, \exists y_1, y_2, u \subseteq_{\mathrm{p}} x\, (\mathrm{Term}(y_1) \wedge \mathrm{Term}(y_2) \wedge (u = \overline{\mathbf{13}} \vee u = \overline{\mathbf{15}})$$
$$\wedge\, y = \mathrm{Subt}(z_1, z_2, y_1)\, u\, \mathrm{Subt}(z_1, z_2, y_2)))).$$

(2) $\mathrm{Form}(x) \to \mathrm{Subf}(z_1, z_2, \mathrm{Neg}(x)) = \mathrm{Neg}(\mathrm{Subf}(z_1, z_2, x))$.

(3) $(\mathrm{Form}(x_1) \wedge \mathrm{Form}(x_2)) \to \mathrm{Subf}(z_1, z_2, \mathrm{Imp}(x_1, x_2))$
$$= \mathrm{Imp}(\mathrm{Subf}(z_1, z_2, x_1), \mathrm{Subf}(z_1, z_2, x_2)).$$

(4) $(\mathrm{Form}(x_1) \wedge \mathrm{Form}(x_2)) \to (\mathrm{Subf}(z_1, z_2, \mathrm{UQ}(x_1, x_2)) = y$
$$\leftrightarrow ((\exists y_1 \subseteq_{\mathrm{p}} x_1\, (\mathrm{Var}(y_1)$$
$$\wedge\, x_1 = y_1\, \overline{\mathbf{15}}\, \overline{\mathbf{14}} \wedge y_1 \neq z_2) \wedge y = \mathrm{UQ}(x_1, \mathrm{Subf}(z_1, z_2, x_2)))$$
$$\vee\, (\exists y_1 \subseteq_{\mathrm{p}} x_1\, (\mathrm{Var}(y_1) \wedge x_1 = y_1\, \overline{\mathbf{15}}\, \overline{\mathbf{14}} \wedge y_1 = z_2)$$
$$\wedge\, y = \mathrm{UQ}(x_1, x_2)))).$$

One can easily prove a lemma corresponding to 4.11 (using the second part of 4.9) which will show that Subf(z_1, z_2, x) is an R_1^+ definable partial function.

We have now shown that most syntactic notions can be naturally defined with R_1^+ formulas of L. We shall, however, require a few more and we list these now giving only the informal interpretations. We hope it is clear that they can all be defined by recursive equations (or suitable characteristic functions can be, in the case of relations) so that the following formulas are all R_1^+. Further, all functions below are total on their obvious domains, the lengths of their values being bounded by a suitable polynomial (i.e., term) in the lengths of their arguments.

4.14. Definitions. (1) For $n \in \omega$, $n \geqslant 1$,

$\mathrm{Subt}_n(z_1, \ldots, z_n, y_1, \ldots, y_n, x) =$ "the result of substituting the terms $z_1, \ldots, z_n$ simultaneously for the variables $y_1, \ldots, y_n$ (respectively) in the term x".

(2) For $n \in \omega$, $n \geqslant 1$,

$\mathrm{Subf}_n(z_1, \ldots, z_n, y_1, \ldots, y_n, x) \overset{\mathrm{def}}{=}$ "the result of substituting the terms $z_1, \ldots, z_n$ simultaneously for the free occurrences of variables $y_1, \ldots, y_n$ (respectively) in the formula x".

(3) $\mathrm{Occ}(z_1, x) \overset{\mathrm{def}}{\Leftrightarrow}$ "the variable z_1 occurs in the term x".

(4) $\mathrm{Nocc}(z_1, x) \overset{\mathrm{def}}{\Leftrightarrow}$ "the variable z_1 does not occur in the term x".

(5) $\mathrm{Freesubf}(z_1, z_2, x) \overset{\mathrm{def}}{\Leftrightarrow}$ "the term z_1 is free for the variable z_2 in the formula x".

(6) $\mathrm{Sent}(x) \overset{\mathrm{def}}{\Leftrightarrow}$ "x is a sentence".

(7) $\mathrm{Val}(x) \overset{\mathrm{def}}{=}$ "the value of the closed term x".

(Thus Val is the unique RUD_1 definable function, with domain $\mathrm{Clterm}(x)$, satisfying the following:

(i) $\mathrm{Z}(x) \rightarrow (\mathrm{Val}(x) = \mathbf{0})$,
(ii) $\mathrm{Clterm}(x) \rightarrow (\mathrm{Val}(\mathrm{Suc}(x)) = \mathrm{Val}(x)')$,
(iii) $(\mathrm{Clterm}(x_1) \wedge \mathrm{Clterm}(x_2)) \rightarrow (\mathrm{Val}(\mathrm{Sum}(x_1, x_2)) = (\mathrm{Val}(x_1) + \mathrm{Val}(x_2)))$,
(iv) $(\mathrm{Clterm}(x_1) \wedge \mathrm{Clterm}(x_2)) \rightarrow (\mathrm{Val}(\mathrm{Prod}(x_1, x_2)) = (\mathrm{Val}(x_1) \cdot \mathrm{Val}(x_2)))$.

That a lemma analogous to 4.11 goes through here follows from the fact that $\forall x\, (|\mathrm{Val}(x)| \leqslant |x|)$ can be proved, by induction on $|x|$, directly from (i)–(iv).)

(8) For $n \in \omega$, $\mathrm{Indax}_n(x) \Leftrightarrow$ "x is an instance of an axiom of $\mathrm{I}\Delta_0 + \Omega_n$" (cf. 4.7(2) and (2) above).

(9) $\mathrm{Subf}(z, y, x) = u \overset{\mathrm{def}}{\Leftrightarrow}$ "z, y are of the form $\overline{16}z_1\overline{16}z_2 \cdots \overline{16}z_t\overline{16}$, $\overline{16}y_1\overline{16}y_2 \cdots \overline{16}y_t\overline{16}$ respectively, where, for $1 \leqslant i \leqslant t$, z_i is a closed term and y_i is a variable, and u is the result of simultaneously substituting z_i for y_i (for all $i = 1, \ldots, t$) in the formula x".

(10) $\mathrm{Reform}(t, x) \overset{\mathrm{def}}{\Leftrightarrow}$ "x is a t-formula", that is, "there is a formula u such that $|u| \leqslant |t|$ and $\mathrm{Subf}(z, y, u) = x$ for some sequence z of closed terms, and some sequence y of variables with $|z|$, $|y| \leqslant |x|$". $(\Leftrightarrow \exists\, |u| \leqslant |t|,\ |z| \leqslant |x|,\ |y| \leqslant |x|$ $\mathrm{Subf}(z, y, u) = x.)$

(*Remark.* We shall only be interested in t-formulas for t standard, that is, in formulas that are substitution instances (by closed terms of arbitrary length) of standard formulas (cf. 3.3, 3.4). However, the fact that this definition is uniform in t will be important in subsequent proofs.)

(11) For $n \in \omega$, $\mathrm{Incvar}_n(x) \overset{\text{def}}{=}$ "the formula obtained when the suffix of every variable occurring in the formula or term x is increased by n" (i.e., changed from $v_{\underbrace{11\cdots1}_{t\text{-times}}}$ to $v_{\underbrace{11\cdots1}_{(t+n)\text{-times}}}$).

(12) For $\phi = \phi(v_1)$ any standard formula with at most the one free variable v_1 and for $m =$ the largest suffix of any variable occurring in ϕ,

$$\mathrm{Rel}_\phi(x) = y \overset{\text{def}}{\Leftrightarrow} \text{"}x \text{ is a formula and } y \text{ is the formula obtained by restricting all quantifiers in } \mathrm{Incvar}_m(x) \text{ to } \phi\text{"}.$$

5. Coding proofs

The formal system we shall be using is taken from [5] and is as follows:

5.1. Axiom schemes. (1) $(\phi \rightarrow (\psi \rightarrow \phi))$.
(2) $((\phi \rightarrow (\psi \rightarrow \chi)) \rightarrow ((\phi \rightarrow \psi) \rightarrow (\phi \rightarrow \chi)))$.
(3) $((\neg\phi \rightarrow \neg\psi) \rightarrow ((\neg\phi \rightarrow \psi) \rightarrow \phi))$.
(4) $(\forall x\, \phi(x) \rightarrow \phi(t))$ (t a term free for the variable x in $\phi(x)$).
(5) $(\forall x\, (\phi \rightarrow \psi) \rightarrow (\phi \rightarrow \forall x\, \psi))$ (x not free in ϕ).
(6) Finitely many quantifier free equality axioms (see [5]).

5.2. Rules. (1) Modus Ponens: ψ following from $(\phi \rightarrow \psi)$ and ϕ.
(2) Generalization: $\forall x\, \phi$ follows from ϕ.

5.3. It is easy to construct R_1^+ formulas $\mathrm{Logax}_i(x)$ (for $i = 1, \ldots, 6$), $\mathrm{MP}(x, y) = z$ and $\mathrm{Gen}(x) = y$ naturally representing "x is an instance of 5.1(i)", "z follows from x and y by modus ponens" and "y follows from x by generaliztion", respectively. For example,

$$\begin{aligned}\mathrm{Logax}_4(x) \overset{\text{def}}{\Leftrightarrow}\ & \exists\, |y_1| \leqslant |x|\ \exists\, |y_2| \leqslant |x|\ \exists\, |y_3| \leqslant |x|\ \exists\, |y_4| \leqslant |x| \\ & (\mathrm{Var}(y_1) \wedge \mathrm{Form}(y_2) \wedge \mathrm{Term}(y_3) \wedge \mathrm{Freesubf}(y_3, y_1, y_2) \\ & \wedge \mathrm{Subf}_1(y_3, y_1, y_2) = y_4 \wedge x = \mathbf{3\,7}\, y_1 y_2\, \mathbf{6}\, y_4\, \mathbf{4}),\end{aligned}$$

$$\mathrm{MP}(x, y) = z \overset{\text{def}}{\Leftrightarrow} y = \mathbf{3}\, x\, \mathbf{6}\, z\, \mathbf{4} \qquad \text{(note a use of 4.9 here)}.$$

We set

$$\mathrm{Logax}(x) \overset{\text{def}}{\Leftrightarrow} \bigvee_{i=1}^{6} \mathrm{Logax}_i(x).$$

Now suppose $A(x)$ is an R_1^+ formula such that $\mathrm{I}\Delta_0 + \Omega_1 \vdash \forall x\,(A(x) \to \mathrm{Sent}(x))$. We define the formula $\mathrm{Proof}_A(x, y)$, "y is a proof of the formula x from the non-logical axioms A", as follows:

$\mathrm{Proof}_A(x, y) \overset{\mathrm{def}}{\Leftrightarrow}$ y is of the form $\mathbf{2}z_1\mathbf{2}z_2\mathbf{2}\cdots\mathbf{2}z_t\mathbf{2}$ where $z_t = x$, and for each $l = 1, \ldots, t$, $\mathrm{Form}(z_l)$ and either (i) $\mathrm{Logax}(z_l)$ or (ii) $A(z_l)$ or (iii) $\exists j, k < l\ \mathrm{MP}(z_j, z_k) = z_l$ or (iv) $\exists j < l\ \mathrm{Gen}(z_j) = z_l$.

Using the techniques of the previous section it is easy to write $\mathrm{Proof}_A(x, y)$ as an R_1^+ formula.

We further define

$$\mathrm{Prov}_A(x) \overset{\mathrm{def}}{\Leftrightarrow} \exists y\, \mathrm{Proof}_A(x, y).$$

(Notice that $\mathrm{Prov}_A(x)$ is not (equivalent to) an R_1^+ formula since we have no hope of bounding that "$\exists y$". The reason that a proof similar to that of 4.4 breaks down is, of course, that the conclusion of the rule modus ponens is *shorter* than the hypotheses.)

The formulas $\mathrm{Reproof}_A(t, x, y)$ and $\mathrm{Reprov}_A(t, x)$, "y is a t-proof of x from A" and "x is t-provable from A" are defined similarly ($\mathrm{Reproof}_A(t, x, y)$ being R_1^+) except that "... for each $l = 1, \ldots, t$, $\mathrm{Form}(z_l)$ and ..." (in the definition of $\mathrm{Proof}_A(x, y)$) is replaced by "... for each $l = 1, \ldots, t$, $\mathrm{Reform}(t, z_l)$ and ...".

5.4. Conventions. Having now completed the rather tedious groundwork and, we hope, convinced the reader that $\mathrm{I}\Delta_0 + \Omega_1$ is a completely adequate theory for naturally arithmetizing syntax, we shall revert to more usual notation, even when working in nonstandard models of $\mathrm{I}\Delta_0 + \Omega_1$, in order to make the proofs of our main theorems clearer. For example, if $M \vDash \mathrm{I}\Delta_0 + \Omega_1$, we shall use expressions like "let $a \in M$ be a formula (term)" meaning $M \vDash \mathrm{Form}(a)$ ($M \vDash \mathrm{Term}(a)$). Thus all standard terms, formulas and proofs are (or have Gödel numbers) in M. Also if $p \in M$ is a proof we shall often write it out in the form $p = \phi_1, \ldots, \phi_t$ (where $t \in M$ and $\phi_i \in M$ is formula for $i = 1, \ldots, t$) even when t and some of the ϕ_i's are nonstandard. Further, if $\phi(x_1, \ldots, x_n)$ is a standard formula and $\tau_1, \ldots, \tau_n \in M$ are terms, then we shall write simply $\phi(\tau_1, \ldots, \tau_n)$ for the unique $b \in M$ such that

$$M \vDash b = \mathrm{Subf}_n(\tau_1, \ldots, \tau_n, \ulcorner x_1 \urcorner, \ldots, \ulcorner x_n \urcorner, \ulcorner \phi(x_1, \ldots, x_n) \urcorner).$$

Finally, if T is a set of standard L-sentences with an obvious R_1^+ definition, $A(x)$ say, (e.g., if T is finite or $\mathrm{I}\Delta_0 + \Omega_n$ (cf. 4.7(2), 4.14(2))), then we shall write $\mathrm{Proof}_T(x, y)$ etc. for $\mathrm{Proof}_A(x, y)$ etc.

6. Some proofs available in models of $\mathrm{I}\Delta_0 + \Omega_1$

Throughout this section M denotes an arbitrary model of $\mathrm{I}\Delta_0 + \Omega_1$.

6.1. Lemma. *Every element of M is named by a unique canonical closed term in*

M. That is, if $a \in M$, then there is a unique $\tilde{a} \in M$ such that $M \models \text{Cancl}\text{term}(\tilde{a})$ and $M \models \text{Val}(\tilde{a}) = a$ (cf. 4.7(6), 4.14(7)). Further,

$$|\tilde{a}| \leq (8 + 2 \cdot B) \cdot (|a| + 1).$$

Proof. The idea of the proof is clear: if "$a = \sum_{i=1}^{t} a_i \cdot B^i$" ($t \in M$, $1 \leq a_i \leq B$ for $i = 0, \ldots, t$), then "$\tilde{a} = (\mathbf{a}_0 + (\mathbf{B} \cdot (\mathbf{a}_1 + (B \cdot \cdots \cdot (\mathbf{B} \cdot \mathbf{a}_t))\cdots)$". More formally, we work in M and use induction on a. If $a = 0$, take $\tilde{a} = \overline{14}$. Suppose $b > 0$ and we have defined $\tilde{a}$ for all $a < b$. Write $b = k + B \cdot t$, where $1 \leq k \leq B$ and $|b| = 1 + |t|$. Set

$$\tilde{b} = 3\,\overline{14}\,\underbrace{\overline{10}\cdots\overline{10}}_{k\ \text{times}}\,\overline{11}\,3\,\overline{14}\,\underbrace{\overline{10}\cdots\overline{10}}_{B\ \text{times}}\,\overline{12}\,\tilde{t}\,44 \qquad (\text{“} = (\mathbf{k} + (\mathbf{B} \cdot \tilde{t}))\text{”}).$$

Clearly the truth of the lemma for t implies $\text{Cancl}\text{term}(\tilde{b})$ and $\text{Val}(\tilde{b}) = b$. Further,

$$|\tilde{b}| \leq (8 + 2 \cdot B) + |\tilde{t}| \leq (8 + 2 \cdot B) + (8 + 2 \cdot B) \cdot (|t| + 1)$$
$$= (8 + 2 \cdot B) \cdot (|b| + 1).$$

The uniqueness is also easily established by induction. □

Henceforth, $\tilde{a}$ will always denote the canonical closed term naming a given by 6.1, for any $a \in M$. Our present aim is to show that if $\phi(x_1, \ldots, x_n)$ is any (standard) R_1^+ formula and $a_1, \ldots, a_n$ are any elements of M such that $M \models \phi(a_1, \ldots, a_n)$, then M contains a k-proof of $\phi(\tilde{a}_1, \ldots, \tilde{a}_n)$ from $I\Delta_0 + \Omega_1$ for some $k \in \omega$ independent of $a_1, \ldots, a_n$.

6.2. Lemma. *There are constants $k_0, k_1, k_2 \in \omega$ such that whenever a, b, $c \in M$ and $M \models ab = c$, then M contains a k_0-proof, p, of the (M-) formula $\tilde{a}\tilde{b} = \tilde{c}$ (from $I\Delta_0 + \Omega_1$) such that $|p| \leq k_1 \cdot |c|^2 + k_2$.*

Proof. The constants k_0, k_1, k_2 are determined by the lengths of some standard proofs and could be easily calculated in advance of the following proof. Since their actual values are unimportant, however, we shall not bother to do this.

We work in M and use induction (on c). Suppose, then, that $ab = c$. If $a = 0$, then $b = c$ so $\tilde{b}$ is the same term as $\tilde{c}$ (by 6.1). Hence the required k_0-proof is obtained by first producing a (standard) proof, from $I\Delta_0$, of $\forall v_1\,(\mathbf{0}v_1 = v_1)$ and then following this with the formulas $(\forall v_1\,(\mathbf{0}v_1 = v_1) \rightarrow (\mathbf{0}\tilde{b} = \tilde{c}))$ (instance of 5.1(4)) and $\mathbf{0}\tilde{b} = \tilde{c}$, i.e., $\tilde{a}\tilde{b} = \tilde{c}$ (by modus ponens). If $a \neq 0$, then $a = ka_1$ and $c = kc_1$ for some k, a_1, c_1 with $1 \leq k \leq B$, $a_1 < a$, $c_1 < c$, $|c| = 1 + |c_1|$. Now $a_1 b = c_1$, so (by our inductive hypothesis) we have a k_0-proof (from $I\Delta_0$), p_1 say, of $\tilde{a}_1\tilde{b} = \tilde{c}_1$ such that $|p_1| \leq k_1 \cdot |c_1|^2 + k_2$. We now add to p_1 a (standard) proof (from $I\Delta_0$) of

$$\forall v_1 \forall v_2 \forall v_3\,(v_1 v_2 = v_3 \rightarrow (\mathbf{k} + (\mathbf{B} \cdot v_1))v_2 = (\mathbf{k} + (\mathbf{B} \cdot v_3)))$$

and then deduce (in finitely many steps, using 5.1(4) and modus ponens)

$$(\mathbf{k} + (\mathbf{B} \cdot \tilde{a}_1))\tilde{b} = (\mathbf{k} + (\mathbf{B} \cdot \tilde{c}_1)),$$

i.e., $\tilde{a}\tilde{b} = \tilde{c}$, which gives the required k_0-proof, p. Further, we clearly have $|p| \leq |p_1| + k_1^* \cdot |c_1| + k_2^*$ (for some constants k_1^*, $k_2^* \in \omega$ which could have been chosen in advance). Hence $|p| \leq k_1 \cdot |c_1|^2 + k_2 + k_1^* \cdot |c_1| + k_2^* \leq k_1 \cdot (|c_1| + 1)^2 + k_2$ (provided k_1, k_2 have been chosen large enough), i.e., $|p| \leq k_1 \cdot |c|^2 + k_2$ as required. □

The following results are proved using arguments similar to the above so we only give outlines of the inductive procedures involved. (The inductions are all possible in M since in every case we establish a bound on the length of the proof that is required to exist, thus making the relevant formula RUD_1—cf. the remarks in 2.5.)

6.3. Lemma. *There are k_0, k_1, $k_2 \in \omega$ such that whenever a, b, $c \in M$, each of the following formulas has a k_0-proof (from $\mathrm{I}\Delta_0 + \Omega_1$) in M, of length bounded by* $(\max(|a|, |b|, |c|)^{k_1} + k_2$.

(1) $(\tilde{a}' = \tilde{b})$, *whenever* $M \models a' = b$.

(2) $|\tilde{a}| = \tilde{b}$, *whenever* $M \models |a| = b$.

(3) $((\tilde{a} + \tilde{b}) = \tilde{c})$, *whenever* $M \models a + b = c$.

(4) $\neg(\tilde{a} = \tilde{b})$, *whenever* $M \models a \neq b$.

(5) $\neg(\tilde{a}\tilde{b} = \tilde{c})$, *whenever* $M \models ab \neq c$.

(6) $((\tilde{a} \cdot \tilde{b}) = \tilde{c})$, *whenever* $M \models a \cdot b = c$.

Proof. We work in M throughout.

(1) Suppose $a' = b$ and $a = (k + (B \cdot t))$ with $1 \leq k \leq B$. If $k < B$, we can use directly the (standard) theorem of $\mathrm{I}\Delta_0 + \Omega_1$

$$\forall v_1 ((\mathbf{k} + v_1)' = (\mathbf{k}' + v_1)).$$

If $k = B$, however, we use

$$\forall v_1 ((\mathbf{B} + (\mathbf{B} \cdot v_1))' = (\mathbf{1} + (\mathbf{B} \cdot v_1')))$$

to convert a k_0-proof, p_1, say, of $\tilde{t}' = \tilde{s}$ (where $s = t'$; note that $|t| < |a|$) into one, p say, of $\tilde{a}' = \tilde{b}$, such that $|p| \leq |p_1| + k_1^* \cdot |s| + k_2^*$ (for some constants k_1^*, $k_2^* \in \omega$).

(2) Suppose $|a| = b$ and write $a = ka_1$ $(1 \leq k \leq B)$ and $b_1' = b$. Then $|a_1| = b_1$ and we can convert a k_0-proof of $|\tilde{a}_1| = \tilde{b}_1$ into one of $|\tilde{a}| = \tilde{b}$ using the $\mathrm{I}\Delta_0 + \Omega_1$-theorem

$$\forall v_1 \forall v_2 (|v_1| = v_2 \rightarrow |(\mathbf{k} + (\mathbf{B} \cdot v_1))| = v_2')$$

together with (1).

(3) Suppose $a + b = c$ and $a = (l_1 + (B \cdot a_1))$, $b = (l_2 + (B \cdot b_1))$, $c = (l_3 + (B \cdot c_1))$ where $1 \leq l_i \leq B$ for $i = 1, 2, 3$. Then $l_1 + l_2 = l_3 + (\varepsilon \cdot B)$ where $\varepsilon = 0$ or 1. If $\varepsilon = 0$, then $a_1 + b_1 = c_1$, and we can convert a k_0-proof of $((\tilde{a}_1 + \tilde{b}_1) = \tilde{c}_1)$

into one of $((\tilde{a}+\tilde{b})=\tilde{c})$ using the $\mathrm{I}\Delta_0+\Omega_1$-theorem

$$\forall v_1\forall v_2\forall v_3\,(((v_1+v_2)=v_3)\rightarrow(((\mathbf{l}_1+(\mathbf{B}\cdot v_1))+(\mathbf{l}_2+(\mathbf{B}\cdot v_2)))\\ =(\mathbf{l}_3+(\mathbf{B}\cdot v_3)))).$$

If $\varepsilon=1$, then $a_1+b_2=c_1$, where $b_2=b_1'$, and we use (1) together with the $\mathrm{I}\Delta_0+\Omega_1$-theorem

$$\forall v_1\forall v_2\forall v_3\forall v_4\,((((v_1+v_4)=v_3)\wedge(v_4=v_2'))\\ \rightarrow(((\mathbf{l}_1+(\mathbf{B}\cdot v_1))+(\mathbf{l}_2+\mathbf{B}\cdot v_2)))=(\mathbf{l}_3+(\mathbf{B}\cdot v_3)))).$$

(4) If $a<b$, then $a+c'=b$, for some c and hence we have, by (1) and (3), a k_0-proof of $((\tilde{a}+\tilde{c}')=\tilde{b})$. A k_0-proof of $\neg(\tilde{a}=\tilde{b})$ can now be obtained by use of the $\mathrm{I}\Delta_0+\Omega_1$-theorem

$$\forall v_1\forall v_2\forall v_3(((v_1+v_2')=v_3)\rightarrow\neg(v_1=v_2)).$$

If $b<a$ the proof is similar.

(5) Suppose $ab=d$ and $d\neq c$. By 6.2 we have a k_0-proof of $\tilde{a}\tilde{b}=\tilde{d}$ and, by (4), one of $\neg(\tilde{d}=\tilde{c})$. Clearly we can now find one of $\neg(\tilde{a}\tilde{b}=\tilde{c})$.

(6) Suppose $a\cdot b=c$. Define b_1, c_1, a_1, k by $b=kb_1$ $(=k+(B\cdot b_1))$, $c_1=a\cdot b_1$ and $a_1=a\cdot k$, where $1\leqslant k\leqslant B$, $|b_1|<|b|$ and $|c_1|<|c|$ (unless $a=0$, in which case the result is easy to show). Notice that $a_1+c_2=c$, where $c_2=B\cdot(a\cdot b_1)$. Now, given a k_0-proof of $((\tilde{a}\cdot\tilde{b}_1)=\tilde{c}_1)$ we add ones of $(\tilde{a}_1=(\tilde{a}\cdot\mathbf{k}))$, $(c_2=(\mathbf{B}\cdot\tilde{c}_1))$ (by repeated use of (3) and the obvious $\mathrm{I}\Delta_0+\Omega_1$-theorem) and $((\tilde{a}_1+\tilde{c}_2)=\tilde{c})$ (using (3)). We can now obtain a k_0-proof of $((\tilde{a}\cdot\tilde{b})=\tilde{c})$ by using the $\mathrm{I}\Delta_0+\Omega_1$-theorem

$$\forall v_1\forall v_2\forall v_3\forall v_4\forall v_5\forall v_6((((v_1\cdot v_2)=v_3)\rightarrow((v_4=(v_1\cdot\mathbf{k}))\rightarrow((v_5=(\mathbf{B}\cdot v_3))\\ \rightarrow((v_6=(v_4+v_5))\rightarrow((v_1\cdot(\mathbf{k}+(B\cdot v_2)))=v_6))))).\quad\square$$

6.4. Theorem. *For any (standard) R_1^+ formula, $\phi(x_1,\ldots,x_n)$ say, of L, there exist k_0, k_1, $k_2\in\omega$ (depending only on ϕ) such that*

($*$) *For all $a_1,\ldots,a_n\in M$, $M\models\phi(a_1,\ldots,a_n)$ implies M contains a k_0-proof from $\mathrm{I}\Delta_0+\Omega_1$, p say, of $\phi(\tilde{a}_1,\ldots,\tilde{a}_n)$ such that*

$$|p|\leqslant(\max|a_1|,\ldots,|a_n|))^{k_1}+k_2.$$

Proof. We use (meta-) induction on ϕ.

If ϕ is of the form $\alpha\beta=\gamma$ or $\neg(\alpha\beta=\gamma)$, where α, β, γ are any (standard) terms of L, then ($*$) clearly follows by repeated use of 6.3 (1), (3), (6) together with 6.2 and 6.3 (5) and the equality axioms.

If ($*$) is true for ϕ, ψ, then it is clearly also true for $(\phi\wedge\psi)$ and for $(\phi\vee\psi)$ (i.e. for $\neg(\phi\rightarrow\neg\psi)$ and for $(\neg\phi\rightarrow\psi)$).

Suppose $\phi(x_1,\ldots,x_n)$ is $\exists\,|y|\leqslant t(|x_1|)\,\psi(y,x_1,\ldots,x_n)$ (t a term of L), i.e., $\exists y\,(|y|\leqslant t(|x_1|)\wedge\psi(y,x_1,\ldots,x_n))$, and that ($*$) holds for $\psi(y,x_1,\ldots,x_n)$. Suppose $a_1,\ldots,a_n\in M$ and $M\models\phi(a_1,\ldots,a_n)$. Choose $a\in M$ such that $M\models$

$|a| \leqslant t(|a_1|)$ and $M \models \psi(a, a_1, \ldots, a_n)$. Now, for suitable k_0, k_1, k_2 (independent of $a, a_1, \ldots, a_n$) we have k_0-proofs, p and q say, of $|\tilde{a}| \leqslant t(|\tilde{a}_1|)$ (using 6.3 (1), (2), (3) and (6), together with the $I\Delta_0 + \Omega_1$-theorem $\forall v_1 \forall v_2((v_1 \leqslant v_2) \leftrightarrow \exists v_3((v_1 + v_3) = v_2)))$ and of $\psi(\tilde{a}, \tilde{a}_1, \ldots, \tilde{a}_n)$ (by the inductive hypothesis) respectively, such that $|p|, |q| \leqslant (\max(|a|, |a_1|, \ldots, |a_n|))^{k_1} + k_2$. However, for some $m_1, m_2 \in \omega$ (depending, only on t, and hence on ϕ) we have $|a| \leqslant |a_1|^{m_1} + m_2$ (since t is a term of L). It now clearly follows that, for suitable $n_0, n_1, n_2 \in \omega$ depending only on ϕ, M contains an n_0-proof (from $I\Delta_0 + \Omega_1$) of $\phi(\tilde{a}_1, \ldots, \tilde{a}_n)$ of length bounded by $(\max |a_1|, \ldots, |a_n|))^{n_1} + n_2$.

If $\phi(x_1, \ldots, x_n)$ is $\exists y \subseteq_p x_1 \psi(y, x_1, \ldots, x_n)$, where $(*)$ holds for ψ, then the proof of $(*)$ for ϕ is similar and is left to the reader.

There only remains the case where ϕ is $\forall y \subseteq_p x_1\, \psi(y, x_1, \ldots, x_n)$ and $(*)$ holds for $\psi(y, x_1, \ldots, x_n)$. We show first, however, that $(*)$ holds for the formula $\forall y \subseteq_i x_1\, \psi(y, x_1, \ldots, x_n)$, where $\forall y \subseteq_i x_1 \cdots$ is the quantifier $\forall y\, (\exists z\, yz = x_1 \rightarrow \cdots)$. As in 6.2 and 6.3, we fix $\boldsymbol{a} = a_1, \ldots, a_n \in M$ and k $(1 \leqslant k \leqslant B)$ and show how to construct for any $d \in M$ a k_0-proof (from $I\Delta_0 + \Omega_1$), p, of $\forall y \subseteq_i \widetilde{dk}\, \phi(y, \tilde{\boldsymbol{a}})$ from a k_0-proof, p_1 say, of $\forall y \subseteq_i \tilde{d}\, \phi(y, \tilde{\boldsymbol{a}})$, assuming that $M \models \forall y \subseteq_i dk\, \phi(y, \boldsymbol{a})$ and that k_0 is sufficiently large. It will be clear that $|p| \leqslant |p_1| + (\max(d, a_1, \ldots, a_n))^{k_1} + k_2$ for some $k_1, k_2 \in \omega$ that could be chosen in advance.

Now the sentence

$$\forall \boldsymbol{x}, z, u\, (u = z\mathbf{k} \rightarrow (\psi(u, \boldsymbol{x}) \rightarrow (\forall y \subseteq_i z\, \psi(y, \boldsymbol{x}) \rightarrow \forall y \subseteq_i u\, \psi(y, \boldsymbol{x}))))$$

is clearly a (standard) theorem of $I\Delta_0 + \Omega_1$. Hence we obtain in M a k_0-proof of (setting $b = dk$)

$$(\tilde{b} = \tilde{d}\mathbf{k} \rightarrow (\psi(\tilde{b}, \tilde{\boldsymbol{a}}) \rightarrow (\forall y \subseteq_i \tilde{d}\, \psi(y, \tilde{\boldsymbol{a}}) \rightarrow \forall y \subseteq_i \tilde{b}\, \psi(y, \tilde{\boldsymbol{a}})))).$$

We now add k_0-proofs of $\tilde{b} = \tilde{a}_1\mathbf{k}$ (using 6.2 — notice that this is exactly the same sentence as $\tilde{b} = \tilde{a}_1\tilde{k}$), and of $\psi(\tilde{b}, \tilde{\boldsymbol{a}})$ (using the original inductive hypothesis since $M \models \forall y \subseteq_i dk\, \psi(y, \boldsymbol{a})$ obviously implies $M \models \psi(b, \boldsymbol{a})$), and p_1, from which we can infer, using Modus Ponens, $\forall y \subseteq_i \tilde{b}\, \psi(y, \tilde{\boldsymbol{a}})$ as required.

Now suppose $M \models \forall y \subseteq_p kd\, \psi(y, \boldsymbol{a})$ $(1 \leqslant k \leqslant B$, d any element of $M)$, so that certainly $M \models \forall y \subseteq_p d\, \psi(y, \boldsymbol{a})$. Hence, by way of another suitable inductive hypothesis, we may suppose M contains a k_0-proof, p_1 say, of $\forall y \subseteq_p \tilde{d}\, \psi(y, \tilde{\boldsymbol{a}})$. To construct a k_0-proof of $\forall y \subseteq_p kd\, \psi(y, \tilde{\boldsymbol{a}})$, we begin with a proof of the (standard) sentence

$$\forall \boldsymbol{x}, z, u\, (u = \mathbf{k}z \rightarrow (\forall y \subseteq_p z\, \psi(y, \boldsymbol{x}) \rightarrow (\psi^*(z, \boldsymbol{x}) \rightarrow \forall y \subseteq_p u\, \psi(y, \boldsymbol{x}))))$$

where $\psi^*(z, \boldsymbol{x})$ is $\forall w \subseteq_i z\, \exists\, |v| \leqslant |z|'\, (v = \mathbf{k}w \wedge \phi(v, \boldsymbol{x}))$. Notice that $M \models \psi^*(d, \boldsymbol{a})$ (since $M \models \forall y \subseteq_p kd\, \psi(y, \boldsymbol{a})$) so that, by the previous cases in this proof we have a k_0-proof, p_2 say, in M (of suitable length) of $\psi^*(\tilde{d}, \tilde{\boldsymbol{a}})$. Since we also have k_0-proofs of $\widetilde{kd} = \mathbf{k}\,\tilde{d}$ (by 6.2) and of $\forall y \subseteq_p d\, \psi(y, \tilde{\boldsymbol{a}})$ (namely p_1), we easily obtain the required proof of $\forall y \subseteq_p \widetilde{kd}\, \psi(y, \tilde{\boldsymbol{a}})$.

We have now, in fact, shown that for any $d, a_1, \ldots, a_n \in M$ such that $M \vDash \forall y \subseteq_p d\, \psi(y, \boldsymbol{a})$, M contains a k_0-proof (from $I\Delta_0 + \Omega_1$) of $\forall y \subseteq_p \bar{d}\, \psi(y, \bar{\boldsymbol{a}})$ of length $\leqslant (\max(d, a_1, \ldots, a_n))^{k_1} + k_2$ (for suitable k_0, k_1, $k_2 \in \omega$ depending only ψ). It remains only to set $d = a_1$ to complete the proof of (*) for ϕ. □

Theorem 6.4 (concerning which, the reader may like to look again at Remark 2.7) will be used many times in subsequent proofs. For the moment we show how it implies that the formula $\mathrm{Prov}_A(x)$ (for suitable A) satisfies the three essential properties of a proof predicate required in Gödel's second incompleteness theorem.

6.5. Theorem. *Let T be any set of L-sentences such that $T \vdash I\Delta_0 + \Omega_1$. Suppose further that $A(x)$ is an R_1^+ formula satisfying $I\Delta_0 + \Omega_1 \vdash \forall x\, (A(x) \rightarrow \mathrm{Sent}(x))$ and $\{\ulcorner \delta \urcorner : \delta \in T\} = \{m \in \omega : \mathbb{N} \vDash A(m)\}$. Let ϕ, ψ be any L-sentences and set $l = \ulcorner \phi \urcorner$, $m = \ulcorner \psi \urcorner$, $n = \ulcorner (\phi \rightarrow \psi) \urcorner$, $r = \ulcorner \mathrm{Prov}_A(\mathbf{l}) \urcorner$. Then we have:*

(1) *$T \vdash \phi$ implies $T \vdash \mathrm{Prov}_A(\mathbf{l})$.*
(2) *$T \vdash (\mathrm{Prov}_A(\mathbf{n}) \rightarrow (\mathrm{Prov}_A(\mathbf{l}) \rightarrow \mathrm{Prov}_A(\mathbf{m})))$.*
(3) *$T \vdash (\mathrm{Prov}_A(\mathbf{l}) \rightarrow \mathrm{Prov}_A(\mathbf{r}))$.*

Proof. (1) If $T \vdash \phi$, then certainly $\mathbb{N} \vDash \mathrm{Prov}_A(l)$, so $M \vDash \mathrm{Prov}_A(l)$ for any $M \vDash T$ (since $\mathrm{Prov}_A(x)$ is clearly preserved in end extensions). Thus $T \vdash \mathrm{Prov}_A(\mathbf{l})$.

(2) Suppose $M \vDash T$, $M \vDash \mathrm{Prov}_A(n)$ and $M \vDash \mathrm{Prov}_A(l)$. Then M contains a proof of $(\phi \rightarrow \psi)$ and one of ϕ. These proofs can clearly be combined to produce one (using modus ponens) of ψ. Hence $M \vDash \mathrm{Prov}_A(m)$. So $M \vDash (\mathrm{Prov}_A(n) \rightarrow (\mathrm{Prov}_A(l) \rightarrow \mathrm{Prov}(m)))$ for an arbitrary $M \vDash T$.

(3) Suppose $M \vDash T$ and $M \vDash \mathrm{Prov}_A(l)$, say $p \in M$ satisfies $M \vDash \mathrm{Proof}_A(l, p)$. Since the L-formula $\mathrm{Proof}_A(\mathbf{l}, y)$ is R_1^+ and $M \vDash I\Delta_0 + \Omega_1$, it follows from 6.4 that M contains a k_0-proof from $I\Delta_0 + \Omega_1$ (for some $k_0 \in \omega$) of the $(M-)$ sentence $\mathrm{Proof}_A(\mathbf{l}, \bar{p})$, and hence one of $\exists y\, \mathrm{Proof}_A(\mathbf{l}, y)$. It follows that $\exists y\, \mathrm{Proof}_A(\mathbf{l}, y)$ has a proof in M using only *standard* axioms from $I\Delta_0 + \Omega_1$ (this is not completely trivial, though we leave it to the reader to check) and hence one using only sentences from T. Therefore $M \vDash \exists y\, \mathrm{Proof}_A(r, y)$. We have shown $M \vDash (\mathrm{Prov}_A(l) \rightarrow \mathrm{Prov}_A(r))$ for any $M \vDash T$. □

6.6. Definitions. Let $A(x)$ be any R_1^+ formula such that $I\Delta_0 + \Omega_1 \vdash \forall x\, (A(x) \rightarrow \mathrm{Sent}(x))$.

(1) $\mathrm{Consis}_A \overset{\text{def}}{\Leftrightarrow} \neg\mathrm{Prov}_A(\mathbf{m})$, where $m = \ulcorner(\mathbf{0} = \mathbf{1})\urcorner$.

(2) $\mathrm{Reconsis}_A(x) \overset{\text{def}}{\Leftrightarrow} \neg\mathrm{Reprov}_A(x, \mathbf{m})$, where $m = \ulcorner(\mathbf{0} = \mathbf{1})\urcorner$.

In particular, if T is a recursive theory and $A(x)$ is a natural R_1^+ formula such that $\{\ulcorner \delta \urcorner : \delta \in T\} = \{n \in \omega : \mathbb{N} \vDash A(n)\}$, then we shall write $\mathrm{Con}(T)$ for Consis_A and $\mathrm{Con}(T, k)$ for $\mathrm{Reconsis}_A(k)$.

Using the usual argument (see [1] for example) we can now deduce from 6.5 the following

6.7. Theorem. *With T, $A(x)$ as in 6.5, $T \nvdash \mathrm{Consis}_A$ (provided T is consistent). In particular for $1 \leqslant n \in \omega$ and any sentence σ consistent with $\mathrm{I}\Delta_0 + \Omega_n$ we have*

$$\mathrm{I}\Delta_0 + \sigma + \Omega_n \nvdash \mathrm{Con}(\mathrm{I}\Delta_0 + \sigma + \Omega_n).$$

As already mentioned we shall improve considerably on 6.7. To this end we require some more results on proofs available in models of $\mathrm{I}\Delta_0 + \Omega_1$, concerning in particular, relativized formulas.

7. Relativizing proofs in models of $\mathrm{I}\Delta_0 + \Omega_1$

7.1. Definition. A (standard) formula $\phi = \phi(v_1)$ of L is called *initial* if

(i) $\mathrm{I}\Delta_0 \vdash \phi(\mathbf{0})$,
(ii) $\mathrm{I}\Delta_0 \vdash \forall x, y\, [(\phi(x) \wedge \phi(y)) \rightarrow (\phi(x') \wedge \phi(x+y) \wedge \phi(x \cdot y))]$, and
(iii) $\mathrm{I}\Delta_0 \vdash \forall x, y\, [(\phi(y) \wedge x \leqslant y) \rightarrow \phi(x)]$.

Thus if ϕ is initial ϕ^M determines an initial segment substructure in every model, M, of $\mathrm{I}\Delta_0$ and hence this substructure will itself be a model of $\mathrm{I}\Delta_0$ (because $\mathrm{I}\Delta_0$ is a Π_1 theory). In other words we have $\mathrm{I}\Delta_0 \vdash \sigma^\phi$ for each $\sigma \in \mathrm{I}\Delta_0$, where σ^ϕ denotes σ relativized to ϕ.

Our proof of 3.1 can now be described informally as follows. For each $n \in \omega$, let ϕ_n be an initial formula such that $\mathrm{I}\Delta_0 \vdash \Omega_n^{\phi_n}$. (The existence of ϕ_n is due to Solovay and we present the construction below.) Now if $\neg\mathrm{Con}(\mathrm{I}\Delta_0 + \Omega_n)$, then $\mathrm{I}\Delta_0 + \Omega_n \vdash \mathbf{0} = \mathbf{1}$, from which it easily follows that $(\mathrm{I}\Delta_0 + \Omega_n)^{\phi_n} \vdash \mathbf{0} = \mathbf{1}$. But by the above remarks $\mathrm{I}\Delta_0 \vdash (\mathrm{I}\Delta_0)^{\phi_n}$, so $\mathrm{I}\Delta_0 \vdash (\mathrm{I}\Delta_0 + \Omega_n)^{\phi_n}$, so $\mathrm{I}\Delta_0 \vdash \mathbf{0} = \mathbf{1}$, i.e., $\neg\mathrm{Con}(\mathrm{I}\Delta_0)$.

Of course, we must now formalize this proof in $\mathrm{I}\Delta_0 + \Omega_1$, so let M once again denote an arbitrary model of $\mathrm{I}\Delta_0 + \Omega_1$. We also fix an initial formula $\phi(v_1)$ and suppose that m is the largest suffix of any variable occurring in ϕ. For ψ a formula or term in M we write ψ^+ for $\mathrm{Incvar}_m^M(\psi)$ and (for ψ a formula) ψ^ϕ for $\mathrm{Rel}_\phi^M(\psi)$ (cf 4.14 (11) and (12)). It is easy to see that $|\psi^+|$, $|\psi^\phi| \leqslant c \cdot |\psi|$ for some $c \in \omega$ (independent of ψ).

7.2. Lemma. *There exist k_0, $k_1 \in \omega$ such that if ψ is any Δ_0 formula in M and $\alpha = \langle x_1, \ldots, x_t \rangle$ is a sequence ($\alpha \in M$) containing each free variable of ψ^+, then there is a proof (from $\mathrm{I}\Delta_0$), p, in M of the sentence*

$$\sigma \overset{\text{def}}{=} \forall x_1 \cdots \forall x_t[(\phi(x_1) \wedge (\phi(x_2) \wedge \cdots \phi(x_t)) \cdots) \rightarrow (\psi^+ \leftrightarrow \psi^\phi)]$$

such that $|p| \leqslant (|\alpha| \cdot |\psi|)^{k_0} + k_1$. Further, if ψ is a k-formula ($k \in \omega$) and $\alpha \in \omega$, then p may be chosen as l-proof for some $l \in \omega$ (depending on ψ and α).

Proof. Throughout this proof $c_0, c_1, \ldots$ denote absolute (standard) constants.

Suppose first that R, P_i, Q_i ($i = 1, 2$) are any formulas in M (with free variables amongst $\boldsymbol{x} = x_1, \ldots, x_t$) and consider the following sentences:

(1) $\forall \boldsymbol{x}\,(R \to (P_1 \leftrightarrow P_1))$,

(2) $\left[\bigwedge_{i=1}^{2} \forall \boldsymbol{x}\,(R \to (P_i \leftrightarrow Q_i))\right] \to \forall \boldsymbol{x}\,(R \to (D(P_1, Q_1) \leftrightarrow D(P_2, Q_2)))$,

(3) $\forall x_j(R \to \phi(x_j)) \to [\tau \to (\forall z, \boldsymbol{x}\,(((\phi(z) \wedge R) \to (P_1 \leftrightarrow Q_1))$
$\to \forall \boldsymbol{x}\,(R \to (U_1(P_1) \leftrightarrow U_2(Q_1)))))]$,

where τ is the sentence $\forall x, y\,((\phi(y) \wedge x \leqslant y) \to \phi(x))$, x_j occurs in $\boldsymbol{x}$, z does not occur free in R, and $U_1(X)$ is either $\forall z \leqslant x_j\, X$ or $\exists z \leqslant x_j\, X$ and $U_2(X)$ is either $\forall z\,(\phi(z) \to (z \leqslant x_j \to X))$ or $\exists z\,(\phi(z) \wedge (z \leqslant x_j \wedge X))$ (respectively). $D(X, Y)$ is $(X \to Y)$ or $\neg X$.

Clearly (1), (2) and (3) are all logically valid schemas and hence have proofs in M consisting of a standard number of steps which are thus of length $\leqslant c_0 \cdot \max(|P_i|, |Q_i|, |R|, |\alpha|)$ ($i = 1, 2$). We now use this remark to prove the lemma by induction (in M) on ψ^+. As before, we do not bother to determine k_0 and k_1 in advance although this could easily be done.

If ψ^+ is quantifier free, the result follows upon taking $P_1 = \psi^+$ and $R = (\phi(x_1) \wedge \cdots \wedge \phi(x_t)) \cdots)$ in (1) and observing that $|P_1|, |R| \leqslant c_1 \cdot |\psi| \cdot |\alpha|$.

Let us now suppose that ψ, and hence ψ^+, is of the form $\psi^+ = \exists x_1\,(x_1 \leqslant x_j \wedge \psi_0^+(x_1, \ldots, x_t))$ (where $x_1 \neq x_j$) and that there is a proof p_0 (from $\mathrm{I}\Delta_0$) in M of $\sigma_0 = \forall x_1, \ldots, x_t((\phi(x_1) \wedge \cdots \wedge \phi(x_t) \cdots) \to (\psi_0^+ \leftrightarrow \psi_0^\phi))$ such that $|p_0| \leqslant (|\alpha| \cdot |\psi_0|)^{k_0} + k_1$. We are looking for a proof (from $\mathrm{I}\Delta_0$) of suitable length of

$$\forall x_2, \ldots, x_t\,((\phi(x_2) \wedge \cdots \wedge \phi(x_t)) \cdots) \to (\psi^+ \leftrightarrow \psi^\phi)).$$

We now set $R(x_2, \ldots, x_t) = (\phi(x_2) \wedge (\phi(x_3) \wedge \cdots \wedge \phi(x_t)) \cdots)$, $P_1 = \psi_0^+$, $Q_1 = \psi_0^\phi$, $\boldsymbol{x} = x_2, \ldots, x_t$, $z = x_1$, $U_1(P_1) = \exists z \leqslant x_j\, P_1$ (which is ψ^+) and $U_2(Q_1) = \exists z\,(\phi(z) \wedge (z \leqslant x_j \wedge Q_1))$ (which is ψ^ϕ), and let q be a proof in M of the corresponding sentence (3), so that $|q| \leqslant c_2 \cdot |\psi_0| \cdot |\alpha|$. Now it is easy to see that there is a proof q_0 in M of $\forall \boldsymbol{x}\,(R \to \phi(x_j))$ (as x_j occurs amongst $x_2, \ldots, x_t$) such that $|q_0| \leqslant c_3 \cdot |\alpha|^2$. Further, there is a standard proof (from $\mathrm{I}\Delta_0$), q_1 say, of τ, since ϕ is initial. Thus, concatenating the proofs p_0, q, q_0, q_1 and using Modus Ponens three times (notice that σ_0 is precisely the formula $\forall z\, \forall \boldsymbol{x}\,((\phi(z) \wedge R) \to (P_1 \leftrightarrow Q_1))$ we obtain a proof p of

$$\forall x_2, \ldots, x_t((\phi(x_2) \wedge \cdots \wedge \phi(x_t)) \cdots) \to (\psi^+ \leftrightarrow \psi^\phi)),$$

such that

$$\begin{aligned} |p| &\leqslant (|\alpha| \cdot |\psi_0|)^{k_0} + k_1 + c_2 \cdot |\psi_0| \cdot |\alpha| + c_3 \cdot |\alpha|^2 + c_4 \\ &\leqslant (|\alpha| \cdot (|\psi_0| + 1))^{k_0} + k_1 \leqslant (|\alpha| \cdot |\psi|)^{k_0} + k_1 \end{aligned}$$

(provided k_0, k_1 have been chosen suitably), as required.

The other induction steps follow similarly from (2) or (3) and are left to the reader.

For the last part of the lemma we merely observe that if ψ is a k-formula, then the proof constructed above contains only l-formulas for some $l \in \omega$. □

7.3. Corollary. *There are k_0, $k_1 \in \omega$ such that if $\tau \in M$ is a sentence of the form $\forall x_1 \forall x_2 \cdots \forall x_t\, \psi(x_1, \ldots, x_t)$, where ψ is a Δ_0 formula in M, then there is a proof (from $\mathrm{I}\Delta_0 + \tau$), p, in M of τ^Φ such that $|p| \leqslant |\tau|^{k_0} + k_1$. (In particular, if τ is an axiom of $\mathrm{I}\Delta_0$, then τ is of the above form and so p is a proof from $\mathrm{I}\Delta_0$ alone.)*

Further, if τ is a k-formula ($k \in \omega$), then p may be chosen an l-proof from some $l \in \omega$.

Proof. Using 7.2, there is a proof (from $\mathrm{I}\Delta_0$), p_0, in M of

$$\forall x_1, \ldots, x_t\, ((\phi(x_1) \wedge \cdots \wedge \phi(x_t)) \cdots) \rightarrow (\psi^+ \leftrightarrow \psi^\Phi))$$

such that $|p_0| \leqslant |\tau|^{c_0} + c_1$ (for some c_0, $c_1 \in \omega$; note that we may suppose that $|\alpha| \leqslant |\tau|$), and from this sentence together with τ^+ we can clearly deduce, in a standard number of steps the sentence

$$\theta = \forall x_1, \ldots, x_t\, ((\phi(x_1) \wedge \cdots \wedge \phi(x_t)) \cdots) \rightarrow \psi^\Phi).$$

Now it is easy to show, by induction in M, that there are c_2, $c_3 \in \omega$ and a proof (in the predicate calculus alone), q, in M of the sentence

$$\theta \rightarrow \forall x_1(\phi(x_1) \rightarrow \forall x_2(\phi(x_2) \rightarrow \cdots \rightarrow \forall x_t(\phi(x_t) \rightarrow \psi^\Phi) \cdots),$$

i.e., of $\theta \rightarrow \tau^\Phi$, such that $|q| \leqslant |\psi|^{c_2} + c_3$. Thus using Modus Ponens we obtain a proof (from $\mathrm{I}\Delta_0 + \tau^+$), p_1, of the required length of τ^Φ. However, for any sentence $\sigma \in M$, there is a proof in M (of length $\leqslant |\sigma|^{c_4} + c_5$, for some constants c_4, $c_5 \in \omega$) from σ of the sentence σ^+ (by induction in M on σ) and hence p_1 can easily be modified to obtain the required proof from $\mathrm{I}\Delta_0 + \tau$ of τ^Φ.

The last part of the corollary is again clear by inspecting the above proof. □

7.4. Lemma. *There are k_0, $k_1 \in \omega$ such that if $\psi \in M$ is a logical axiom (i.e. $M \models \mathrm{Logax}(\psi)$, cf. 6.5.3), then there is a proof p in M of the formula $\sigma \overset{\mathrm{def}}{=} (\phi(x_1) \wedge \cdots \wedge \phi(x_t)) \cdots) \rightarrow \psi^\Phi$ such that $|p| \leqslant |\sigma|^{k_0} + k_1$, where the free variables of ψ^+ are amongst $x_1, \ldots, x_t$. Further, if σ is a k-formula ($k \in \omega$), then p may be chosen an l-proof for some $l \in \omega$.*

Proof. If ψ is an instance of 5.1 (1), (2), (3) or (6), then ψ^Φ is an instance of the same scheme ($\psi^+ = \psi^\Phi$ in the case of (6)) so the result follows by a use of 5.1 (1) and Modus Ponens. Also if ψ is an instance of 5.1(5), then ψ^Φ is provable in a standard number of steps and the result again follows as above. Thus it only remains to construct a proof of a formula of the form

$$(*) \qquad R \rightarrow [\forall x^+(\phi(x^+) \rightarrow \theta^\Phi(x^+)) \rightarrow \theta(t)^\Phi],$$

(notice $\theta(x)^\Phi = \theta^\Phi(x^+)$), where t is a term free for x in θ, the free variables of

θ^+, t^+ are amongst x^+, $x_1, \ldots, x_t$, and R is the formula $(\phi(x_1) \wedge \cdots \wedge \phi(x_t)) \cdots)$. Now clearly t^+ is free for x^+ in the formula $\phi(x^+) \to \theta^\phi(x^+)$ and $\theta(t)^\phi = \theta^\phi(t^+)$. Hence the formula $(\forall x^+(\phi(x^+) \to \theta^\phi(x^+) \to (\phi(t^+) \to \theta(t)^\phi))$ is an instance of 5.1(6). Thus to construct the required proof of (*) it clearly suffices to find one, q say, of $R \to \phi(t^+)$ such that $q \leqslant (\max(|R|, |t^+|)^{c_0} + c_1$ for some c_0, $c_1 \in \omega$. However, such a q is easily constructed by iterating conditions (i) and (ii) in Definition 7.1 (i.e., using induction in M on t).

The last part of the lemma is again easily checked. □

7.5. Corollary. *Let χ be a (standard) sentence of L such that $\mathrm{I}\Delta_0 \vdash \chi^\phi$ (where ϕ is, as above, an initial formula). Then $M \models \mathrm{Con}(\mathrm{I}\Delta_0) \to \mathrm{Con}(\mathrm{I}\Delta_0 + \chi)$. Moreover, $M \models \mathrm{Con}(\mathrm{I}\Delta_0 + \sigma) \to \mathrm{Con}(\mathrm{I}\Delta_0 + \chi + \sigma)$ for any (standard) Π_1 sentence σ. Further, for all $k \in \omega$ there exists $l \in \omega$ such that $M \models \mathrm{Con}(\mathrm{I}\Delta_0, l) \to \mathrm{Con}(\mathrm{I}\Delta_0 + \chi, k)$ and $M \models \mathrm{Con}(\mathrm{I}\Delta_0 + \sigma, l) \to \mathrm{Con}(\mathrm{I}\Delta_0 + \chi + \sigma, k)$ for σ Π_1.*

Proof. We prove the second assertion, the others being left to the reader.

Suppose $M \models \neg\mathrm{Con}(\mathrm{I}\Delta_0 + \chi + \sigma)$, say $p = \psi_1, \psi_2, \ldots, \psi_u$ is a proof in M from the non-logical axioms $\mathrm{I}\Delta_0 + \chi + \sigma$ and ψ_u is the sentence $\mathbf{0} = \mathbf{1}$. We show how to construct a proof, q, of $\mathbf{0} = \mathbf{1}$ from $\mathrm{I}\Delta_0 + \sigma$ alone. (It will be clear that if p is a k-proof ($k \in \omega$), then the q we construct is an l-proof for some $l \in \omega$.)

Suppose that all the variables (free and bound) occurring in some ψ_α occur amongst $x_1, \ldots, x_t$, and let $R = R(x_1^+, \ldots, x_t^+) = (\phi(x_1^+) \wedge \cdots \wedge \phi(x_t^+)) \cdots)$. For $\alpha \leqslant u$, set $p_\alpha = \psi_1, \psi_2, \ldots, \psi_\alpha$. We show, by induction (in M) on α, that there is a proof from $\mathrm{I}\Delta_0 + \sigma$, q_α say, such that q_α contains all the formulas $R \to \psi_1^\phi, \ldots, R \to \psi_\alpha^\phi$ and such that $|q_\alpha| \leqslant (\max |p_\alpha|, |R|)^{k_0} + k_1$ for some fixed k_0, $k_1 \in \omega$ that could be determined in advance.

Suppose then that q_α has been constructed for some $\alpha < u$. If $\psi_{\alpha+1}$ is either a logical axiom or an axiom of $\mathrm{I}\Delta_0 + \sigma$, then we simply add to q_α the proof (from $\mathrm{I}\Delta_0 + \sigma$) of $R \to \psi_{\alpha+1}^\phi$ given by 7.3 or 7.4 (together with a use of 5.1(1) and Modus Ponens in the latter case). If $\psi_{\alpha+1}$ is χ, then we add to q_α a standard proof (from $\mathrm{I}\Delta_0$) of χ^ϕ, again followed by $\chi^\phi \to (R \to \chi^\phi)$ (Axiom 5.1(1)) and $R \to \chi^\phi$ (Modus Ponens). Now suppose that $\psi_{\alpha+1}$ follows by Modus Ponens from previous formulas, i.e., that for some β, $\gamma < \alpha + 1$, ψ_β is $\psi_\gamma \to \psi_{\alpha+1}$. Now q_α contains the formulas $R \to \psi_\gamma^\phi$ and $R \to \psi_\beta^\phi$, i.e., $R \to (\psi_\gamma^\phi \to \psi_{\alpha+1}^\phi)$, so by using a suitable instance of 5.1(2) and Modus Ponens (twice) we obtain the required proof, $q_{\alpha+1}$, of $R \to \psi_{\alpha+1}^\phi$. Finally, if $\beta < \alpha + 1$ and $\psi_{\alpha+1}$ follows from ψ_β by 5.2(2), then $\psi_{\alpha+1}$ is $\forall x\, \psi_\beta$, q_α contains the formula $R \to \psi_\beta^\phi$, x^+ occurs amongst $x_1^+, \ldots, x_t^+$, and the formula to be proved is $R \to \forall x^+(\phi(x^+) \to \psi_\beta^\phi)$. Let R' be the conjunction of those $\phi(x_i)$ with $i \leqslant t$ and $x_i \neq x^+$. Then it is easy to construct a proof of $R \to R'$ and (using the formula $R \to \psi_\beta^\phi$) one of $R' \to (\phi(x^+) \to \psi_\beta^\phi)$, and hence (using 5.1(5)) one of $R' \to \forall x^+(\phi(x^+) \to \psi_\beta^\phi)$. The formula $R \to \forall x^+(\phi(x^+) \to \psi_\beta^\phi)$ can now be deduced in a further standard number of obvious steps giving the required proof $q_{\alpha+1}$.

The induction is now complete, giving us a proof from $I\Delta_0 + \sigma$ of the formula $R(x_1^+, \ldots, x_t^+) \to \mathbf{0} = \mathbf{1}$. Now using 5.2(2) and 5.1(4), we obtain one of $R(\mathbf{0}, \mathbf{0}, \ldots, \mathbf{0}) \to \mathbf{0} = \mathbf{1}$. But using (i) in Definition 7.1, it is clear that M contains a proof of $R(\mathbf{0}, \mathbf{0}, \ldots, \mathbf{0})$ (from $I\Delta_0$ alone), and hence one of $\mathbf{0} = \mathbf{1}$ (from $I\Delta_0 + \sigma$). Thus $M \models \neg\mathrm{Con}(I\Delta_0 + \sigma)$. $\square$

We remark here, without proof, that by an unpublished result of Wilkie there is a formula $\delta(x)$ such that for Robinson's Arithmetic Q,

$$\begin{aligned} &\mathrm{Q} \vdash \delta(0) \wedge \forall x, y, z\, ((\delta(x) \wedge \delta(y) \wedge z \leqslant y) \\ &\quad \to (\delta(z) \wedge \delta(x+1) \wedge \delta(x+y) \wedge \delta(x \cdot y))), \end{aligned}$$

and furthermore for any axiom θ of $I\Delta_0$, $\mathrm{Q} \vdash \theta^\delta$. Using this it is possible to show in a similar fashion to 7.5 that

$$I\Delta_0 + \Omega_1 \vdash \mathrm{Con}(\mathrm{Q}) \leftrightarrow \mathrm{Con}(I\Delta_0).$$

We now give two applications of 7.5. For the first we need the following lemma. (The statement in parentheses in this lemma will only be required in Section 8.)

7.6. Lemma. *Suppose that $\phi(\boldsymbol{x})$ is a Δ_0 (or a RUD_1) formula of L, that T is a Π_1 theory containing $I\Delta_0$ and that $T + \exp \vdash \forall \boldsymbol{x}\, \phi(\boldsymbol{x})$. Then for some $n \in \omega$, the sentence $\forall y\, \forall \boldsymbol{x} \leqslant y\, (\exists z\, z = B_{n+1}(y) \to \phi(\boldsymbol{x}))$ is provable from T, where, for $j \in \omega$, we define $B_j(x) \overset{\text{def}}{=} e_j(\mathbf{B}, \ldots, \mathbf{B}, x)$.* (Recall from 2.1 that B is the base for our coding of sequences.)

Proof. First suppose that $\phi(\boldsymbol{x})$ is Δ_0 and that there is no such n. Let c be a new constant symbol and consider the theory

$$T' = T + \{\exists z\, z = B_{n+1}(c) \wedge \exists \boldsymbol{x} \leqslant c\, \neg\phi(\boldsymbol{x}) : n \in \omega\}.$$

Clearly our supposition implies that T' is consistent, so let $M \models T'$. Then for all $n \in \omega$, $M \models \exists z\, z = B_{n+1}(c)$, so that the substructure, M^* say, of M with domain $\{a \in M : \exists n \in \omega, M \models a \leqslant B_{n+1}(c)\}$ is clearly an initial segment of M satisfying exp and also satisfying T, since T is a Π_1 theory. Further, since $M \models \exists \boldsymbol{x} \leqslant c\, \neg\phi(\boldsymbol{x})$ and Δ_0-formulas are preserved down in initial segments we have $M^* \models \exists \boldsymbol{x}\, \neg\phi(\boldsymbol{x})$. But this contradicts the assumption that $T + \exp \vdash \forall \boldsymbol{x}\, \phi(\boldsymbol{x})$.

For $\phi(\boldsymbol{x})$ a RUD_1 formula the same proof works because RUD_1 formulas are also preserved down in initial segments of models of $I\Delta_0$ providing the smaller model satisfies Ω_1 (note that $M^* \models \exp$ in the above proof). This follows from the fact that for any term, t, of L,

$$I\Delta_0 + \Omega_1 \vdash \forall y\, \exists z\, \forall x\, (|x| \leqslant t(|y|) \to x \leqslant z). \qquad \text{(cf. 2.5.)} \quad \square$$

7.7. Corollary. *Suppose* $\mathrm{I}\Delta_0 + \exp \vdash \forall \boldsymbol{x}\, \psi(\boldsymbol{x})$ *with* $\psi(\boldsymbol{x})$ Δ_0. *Then*

(i) *There is an initial formula* $\phi(x)$ *such that* $\mathrm{I}\Delta_0 \vdash (\forall \boldsymbol{x}\, \psi(\boldsymbol{x}))^{\phi}$.

(ii) $\mathrm{I}\Delta_0 + \Omega_1 + \mathrm{Con}(\mathrm{I}\Delta_0) \vdash \mathrm{Con}(\mathrm{I}\Delta_0 + \forall \boldsymbol{x}\, \psi(\boldsymbol{x}))$.

(iii) $\mathrm{I}\Delta_0 + \Omega_1 + \mathrm{Con}(\mathrm{I}\Delta_0) \vdash \mathrm{Con}(\mathrm{I}\Delta_0 + \mathrm{Pr})$, *where* Pr *is the sentence expressing the unboundedness of the set of prime numbers* (cf. the example 1.2).

Proof. (i) Define sequences of formulas $Q_i(x)$, $P_i(x)$ (for $i \in \omega$) by

$$Q_0(x) \stackrel{\text{def}}{=} P_0(x) \stackrel{\text{def}}{=} (x = x),$$

$$Q_{i+1}(x) \stackrel{\text{def}}{=} \exists y\, (y = 2^x \wedge P_i(y)),$$

$$P_{i+1}(x) \stackrel{\text{def}}{=} \forall z\, (Q_i(z) \rightarrow Q_i(z \cdot x)).$$

Then it is easy to show that the $P_i(x)$ are initial formulas and that

$$\mathrm{I}\Delta_0 \vdash \forall x, y\, ((Q_i(x) \wedge Q_i(y)) \rightarrow Q_i(x + y)), \quad \text{and}$$
$$\mathrm{I}\Delta_0 \vdash \forall x\, (P_i(x) \rightarrow \exists y\, (y = B_i(x))).$$

Now using 7.6 and our hypothesis let n be such that

$$\mathrm{I}\Delta_0 \vdash \forall y\, \forall \boldsymbol{x} \leqslant y\, (\exists z\, z = B_{n+1}(y) \rightarrow \psi(\boldsymbol{x}))$$

and take $\phi(x)$ to be $P_{n+1}(x)$. Then we obtain

$$\mathrm{I}\Delta_0 \vdash \forall y\, (\phi(y) \rightarrow \forall \boldsymbol{x} \leqslant y\, \psi(\boldsymbol{x}))$$

and hence $\mathrm{I}\Delta_0 \vdash (\forall \boldsymbol{x}\, \psi(\boldsymbol{x}))^{\phi}$ by 7.2 (taking M there to be the standard model).

(ii) This is immediate from (i) and 7.5.

(iii) This part follows from (ii). For although the statement that there exists unboundedly many primes is not of the required Π_1 form it is a consequence of Bertrand's 'conjecture', i.e.

$$\forall x > 0\, \exists \text{ prime } p,\, x \leqslant p \leqslant 2x,$$

and this theorem *is* Π_1 and provable from $\mathrm{I}\Delta_0 + \exp$. □

For our second application of 7.5 we must first give Solovay's definition [7] of formulas $J_n(x)$ satisfying $\mathrm{I}\Delta_0 \vdash \Omega_n^{J_n}$ for all $n \in \omega$.

First, let us write $\eta_n(y, x)$ for $e_n(y, |y|, \ldots, |y|^{(n)}, |x|^{(n+1)})$.

7.8. Definition.

(1) $J_0(x) \stackrel{\text{def}}{\Leftrightarrow} x = x$.

(2) $J_{n+1}(x) \stackrel{\text{def}}{\Leftrightarrow} J_n(x) \wedge \forall y\, (J_n(y) \rightarrow \exists z\, (z = \eta_n(y, x) \wedge J_n(z)))$.

Now it is easy to show that

$$\mathrm{I}\Delta_0 \vdash \forall x, y, z\, (z = \eta_n(y, x) \rightarrow \forall v \leqslant y\, \forall u \leqslant x\, \exists t\, (t = \eta_n(v, u) \wedge t \leqslant z)),$$

and hence, by induction on n, we have

7.9. Lemma. *For all $n \in \omega$,*

$$\mathrm{I}\Delta_0 \vdash \forall x, y\, ((J_n(y) \wedge x \leqslant y) \rightarrow J_n(x)). \quad \square$$

For $n,\ k \in \omega,\ k \geqslant 1$, define $\eta_n^{(1)}(y, x) = \eta_n(y, x)$ and $\eta_n^{(k+1)}(y, x) = \eta_n^{(k)}(\eta_n(y, x), x)$. Then by routine calculations it is easy to check that for all n, $r \in \omega$, there are k, $C \in \omega$ such that

7.10. $\mathrm{I}\Delta_0 \vdash \forall y \geqslant \mathbf{C}\, \forall x, z\, (z = \eta_n^{(k)}(y, x) \rightarrow \exists u\, (u = \omega_{n+1}(x) \wedge u \leqslant z),$

and

7.11. $\mathrm{I}\Delta_0 \vdash \forall u\, (u = \omega_n^{(r)}(y) \rightarrow \exists z\, (z = \eta_n(y, \mathbf{r}) \wedge z \leqslant u))$, where $\omega_n^{(1)}(y) = \omega_n(y)$ and $\omega_n^{(k+1)}(y) = \omega_n^{(k)}(\omega_n(y))$.

(For example, for $n = 0$, we have for all x and all $y \geqslant B$, $\eta_0^{(3)}(y, x) = y^{|x|^3} \geqslant x^{|x|} = \omega_1(x)$, which establishes 7.10 in this case; for 7.11 with $n = 1$, we have $\omega_n^{(r)}(y) \geqslant y^{|y|^r} \geqslant y^{|y|^{||r||}} = \eta_1(y, r)$.)

7.12. Lemma. *For all $n \in \omega$,*
(1) $\mathrm{I}\Delta_0 \vdash \forall x\, (J_n(x) \rightarrow \exists z\, (z = \omega_n(x) \wedge J_n(z)))$,
(2) $\forall r \in \omega$, $\mathrm{I}\Delta_0 \vdash J_n(\mathbf{r})$, *and*
(3) $\mathrm{I}\Delta_0 \vdash \forall x, y\, ((J_n(x) \wedge J_n(y)) \rightarrow (J_n(x + y) \wedge J_n(x \cdot y) \wedge J_n(x')))$.

Proof. We use induction on n to prove (1), (2) and (3) simultaneously.
The case $n = 0$ being clear, suppose the lemma holds for some n. Working in $\mathrm{I}\Delta_0$, notice that iterating 7.8(2) yields, for any $k \in \omega$, $k \geqslant 1$,

$$\forall x\, [J_{n+1}(x) \rightarrow \forall y\, (J_n(y) \rightarrow \exists z\, (z = \eta_n^{(k)}(y, x) \wedge J_n(z)))].$$

Now let k, C be as in 7.10. Then $J_n(\mathbf{C})$ (by the inductive hypothesis) and so

$$\forall x\, [J_{n+1}(x) \rightarrow \exists z\, (z = \eta_n^{(k)}(\mathbf{C}, x) \wedge J_n(z))],$$

and hence by 7.10 and 7.9, $\forall x\, [J_{n+1}(x) \rightarrow \exists z\, (z = \omega_{n+1}(x) \wedge J_n(z))]$, which establishes (1).
For (2), suppose $r \in \omega$ and $J_n(y)$. Then by the inductive hypothesis it follows that for all $k \in \omega$, $\exists u\, (u = \omega_n^{(k)}(y) \wedge J_n(u))$. Taking $k = r$ here, we obtain, by 7.11 and 7.9, $\exists z\, (z = \eta_n(y, \mathbf{r}) \wedge J_n(z))$. Thus we have shown that $\forall y\, (J_n(y) \rightarrow \exists z\, (z = \eta_n(y, \mathbf{r}) \wedge J_n(z))$. Since we have $J_n(\mathbf{r})$ (by the inductive hypothesis), $J_{n+1}(\mathbf{r})$ follows which establishes (2).
(3) follows immediately from (1), (2) and 7.7 since for some $D \in \omega$,

$$\mathrm{I}\Delta_0 \vdash \forall x, y \geqslant \mathbf{D}\, \forall z\, (z = \omega_{n+1}(\max(x, y)) \rightarrow z \geqslant x \cdot y,\ x + y,\ x'). \quad \square$$

Now using 7.9, 7.12 and 7.2 for the Δ_0 formula $\omega_n(x)=z$ we have:

7.13. Corollary. *For all $n \in \omega$, $J_n(x)$ is an initial formula such that* $\mathrm{I}\Delta_0 \vdash \Omega_n^{J_n}$.

The following theorem which includes 3.1 and 3.2, is an immediate consequence of 7.5, 7.13 and 6.7:

7.14. Theorem. *Let $n \in \omega$ and σ be any Π_1 sentence. Suppose T is either* $\mathrm{I}\Delta_0$ *or* $\mathrm{I}\Delta_0+\sigma$. *Then*

(1) $\mathrm{I}\Delta_0+\Omega_1+\mathrm{Con}(T) \vdash \mathrm{Con}(T+\Omega_n)$.

(2) *For all $k \in \omega$, there exists $l \in \omega$ such that* $\mathrm{I}\Delta_0+\Omega_1+\mathrm{Con}(T,\mathbf{l}) \vdash \mathrm{Con}(T+\Omega_n,\mathbf{k})$.

(3) $T+\Omega_n \nvdash \mathrm{Con}(T)$.

7.15. Remarks. The proof of the main result of this section relies heavily on the existence, for each $n \in \omega$, of a formula defining, in each model of $\mathrm{I}\Delta_0$, an initial segment closed under the function ω_n. It is therefore natural to ask whether there are formulas with even stronger closure properties. Now in [3] it is shown that there is no such formula 'closed' under exponentiation. However, in a forthcoming paper the present authors will improve this by showing that no formula defines, in every model of $\mathrm{I}\Delta_0$, an initial segment closed under *all* the ω_n. The precise result is as follows:

Proposition. *Let $M \models \mathrm{I}\Delta_0+\exp$, M non-standard and countable. Then for each $n \in \omega$ there is $k \in \omega$ such that for all $a \in M\setminus\omega$ there is an initial segment $I \subseteq M$ such that*

(1) *$a \in I$, $2^a \notin I$ and I is closed under multiplication (so that $I \models \mathrm{I}\Delta_0$).*

(2) *For any Σ_n formula $\phi(x)$ (possibly with parameters from I), if $I \models \phi(0) \wedge \forall x\,(\phi(x) \rightarrow \phi(x'))$, then $I \models \phi(|a|^{(k)})$.*

To see that this proposition justifies the above remarks first note that for each $r \in \omega$, there is clearly $n=n(r) \in \omega$ with the property that whenever $\phi(x)$ is a Σ_r formula then for *all* $k \in \omega$, the formula

$$\phi^{(k)}(x) \stackrel{\text{def}}{=} \phi(x) \wedge \exists y\,(\phi(y) \wedge y=e_k(x))$$

is a Σ_n formula.

Now suppose $\psi(x)$ defines an initial segment closed under *all* the functions ω_n in every model of $\mathrm{I}\Delta_0$. Say $\psi(x)$ is a Σ_r formula. Now by the proposition, with $n=n(r)$ (and M any non-standard countable model of $\mathrm{I}\Delta_0+\exp$), we can find $I \models \mathrm{I}\Delta_0$ ($I \subseteq_e M$), $k \in \omega$ and $a \in M\setminus\omega$ such that $a \in I$, $2^a \notin I$ and such that (2) holds for all the formulas $\psi^{(p)}(x)$ ($p \in \omega$).

Now it is not hard to show (see the next section) that an initial segment, J, of any model, K, of $\mathrm{I}\Delta_0$ is closed under all the functions ω_n if and only if for all

$p \in \omega$, and for all $b \in J$,

$$J \models \exists x\,(x = B_p(b)) \rightarrow \exists x\,(x = B_p(b+1)).$$

It follows that for all $p \in \omega$,

$$I \models \psi^{(p)}(0) \wedge \forall x\,(\psi^p(x) \rightarrow \psi^p(x'))$$

and hence $I \models \psi^{(p)}(|a|^{(k)})$. Taking $p = k+2$ here we obtain

$$I \models \exists y\,(\psi(y) \wedge y = e_{k+2}(|a|^{(k)})).$$

But $M \models e_{k+2}(|a|^{(k)}) \geqslant 2^a$, so in particular $2^a \in I$—a contradiction.

Perhaps we should now point out that these remarks do not imply that Theorem 7.14 is best possible! Indeed there is a function f, with a Δ_0-definable graph, such that

$$\mathrm{I}\Delta_0 + \Omega_1 + \mathrm{Con}(\mathrm{I}\Delta_0) \vdash \mathrm{Con}(\mathrm{I}\Delta_0 + \forall x\,\exists y\,(f(x) = y))$$

and such that f 'majorizes' all the ω_n, i.e., whenever $M \models \mathrm{I}\Delta_0 + \forall x\,\exists y\,(f(x) = y)$, then for all $n \in \omega$, $M \models \Omega_n$. However, the construction of such an f is rather artificial and will not be needed in the sequel, so we do not present it here.

8. Consistency versus exponentiation

In this section we shall prove Theorems 3.3, 4, 5 and 6 and present some related results. We first require the following

8.1. Lemma. *Let $M \models \mathrm{I}\Delta_0 + \Omega_1$. Then for each $n \in \omega$ there is a $k \in \omega$ such that for each $a \in M$ the sentence $\exists z\,(z = B_n(\bar{a}))$ has a k-proof in M from* $\mathrm{I}\Delta_0$. (Recall from 2.1 that B is the base for our coding of sequences. Also, $B_n(\bar{a}) = e_n(\mathbf{B}, \ldots, \mathbf{B}, \bar{a})$ where (cf. 4.7, 6.1) $\bar{a}$ is the unique canonical term of M with value a.)

Proof. Define sequences of formulas $Q_i(x)$, $P_i(x)$ by

$$Q_0(x) \stackrel{\text{def}}{=} (x = x) \stackrel{\text{def}}{=} P_0(x).$$

$$P_{i+1}(y) \stackrel{\text{def}}{=} \forall z\,(Q_i(z) \rightarrow \exists x\,(x = z^{|y|} \wedge Q_i(x))).$$

$$Q_{i+1}(x) \stackrel{\text{def}}{=} \exists y\,(y = \mathbf{B}^x \wedge P_{i+1}(y)).$$

Then, as for the formulas J_i of Section 7, the P_i, Q_i are initial and indeed $\mathrm{I}\Delta_0 \vdash \forall x\,(P_i(x) \rightarrow \exists y\,(y = \omega_1(x) \wedge P_i(y)))$. Furthermore

$(*)$ $\quad \mathrm{I}\Delta_0 \vdash \forall x\,(Q_i(x) \rightarrow \exists y\,(y = B_i(x))).$

Let $a \in M$, say $a = \sum_{i=0}^{t} a_i \cdot B^i$ ($t \in M$, $1 \leqslant a_i \leqslant B$, $i = 0, \ldots, t$), so that $\bar{a} = (\mathbf{a}_0 + (\mathbf{B} \cdot (\mathbf{a}_1 + \mathbf{B} \cdot (\cdots + (\mathbf{B} \cdot \mathbf{a}_t))\cdots)$. Then using the fact that Q_n is initial we can successively produce in M k-proofs (for suitably large k) of

$$Q_n(\mathbf{B}),\quad Q_n(\mathbf{a}_t),\quad Q_n(\mathbf{a}_{t-1}),\quad Q_n(\mathbf{a}_{t-1}+\mathbf{B}\cdot\mathbf{a}_t),\quad Q_n(\mathbf{a}_{t-2}),$$
$$Q_n(\mathbf{a}_{t-1}+\mathbf{B}\cdot(\mathbf{a}_{t-1}+\mathbf{B}\cdot\mathbf{a}_t)),\ldots,Q_n(\tilde{a}).$$

The result now follows using (*) and Modus Ponens. □

We can now prove a version of 3.3.:

8.2. Theorem. *Let τ be any sentence of L and suppose that M is a countable model of* $\mathrm{I}\Delta_0+\Omega_1$ *such that for all $k\in\omega$, $M\vDash\mathrm{Con}(\mathrm{I}\Delta_0+\tau, k)$ (cf. 6.6(2)). Then there is a model M^* such that*

(i) $M\subseteq_+ M^*$ (*cf.* 2.6),

(ii) $M^*\vDash\mathrm{I}\Delta_0+\tau$,

(iii) *for all $n\in\omega$, $a\in M$, $M^*\vDash\exists z\,(z=B_n(a))$.*

Proof. We have $M\vDash\neg\mathrm{Reprov}_{\mathrm{I}\Delta_0+\tau}(k,\mathbf{0}=\mathbf{1})$ for all $k\in\omega$. Now let $\{\sigma_r:r\in\omega\}$ be an enumeration of all those sentences of M which are k-formulas for some $k\in\omega$. (Notice that this set is countable because M is.) Define, externally, an enumeration $\{\sigma_r^*:r\in\omega\}$ such that for all $r\in\omega$, σ_r^* is either σ_r or $\neg\sigma_r$ and

$$M\vDash\neg\mathrm{Reprov}_{\mathrm{I}\Delta_0+\tau}\left(k,\bigwedge_{i=0}^{r}\sigma_i^*\rightarrow\mathbf{0}=\mathbf{1}\right)$$

for all $k\in\omega$. This definition is easily carried out, by induction on r, since it is trivial to check that if k_0, $k_1\in\omega$ and

$$M\vDash\mathrm{Reprov}_{\mathrm{I}\Delta_0+\tau}\left(k_j,\left(\bigwedge_{i=0}^{r}\sigma_i^*\wedge\sigma_{r+1}^{j}\right)\rightarrow\mathbf{0}=\mathbf{1}\right)\quad\text{for } j=0,1,$$

(where $\sigma^0=\sigma$, $\sigma^1=\neg\sigma$), then for some $k_2\in\omega$,

$$M\vDash\mathrm{Reprov}_{\mathrm{I}\Delta_0+\tau}\left(k_2,\bigwedge_{i=0}^{r}\sigma_i^*\rightarrow\mathbf{0}=\mathbf{1}\right).$$

Now let $L(M)$ be the language obtained from L by adding a new constant symbol, c_a, for each $a\in M$ and set

$$T=\{\phi(c_{a_1},\ldots,c_{a_l}):\ \phi(x_1,\ldots,x_l)\in L,\ a_1,\ldots,a_l\in M,$$
$$\phi(\tilde{a}_1,\ldots,\tilde{a}_l)\in\{\sigma_r^*:r\in\omega\}\}.$$

Now T certainly contains, for each sentence σ of $L(M)$, either σ or $\neg\sigma$, because if $\sigma=\phi(c_{a_1},\ldots,c_{a_l})$ $(\phi(x_1,\ldots,x_l)\in L)$, then $\phi(\tilde{a}_1,\ldots,\tilde{a}_l)$ (as a sentence in M) is a k-formula for some $k\in\omega$, and hence occurs in our original enumeration $\{\sigma_r:r\in\omega\}$. Further, if $\sigma\in T$ and $\neg\sigma\in T$, then we would clearly have, for some k, $r\in\omega$,

$$M\vDash\mathrm{Reprov}_{\mathrm{I}\Delta_0+\tau+\Omega_n}\left(k,\bigwedge_{i=0}^{r}\sigma_i^*\rightarrow\mathbf{0}=\mathbf{1}\right),$$

a contradiction. Thus T is a complete consistent theory. Notice also that if $\phi(x_1, \ldots, x_l) \in L$, $k \in \omega$ and $M \vDash \mathrm{Reprov}_{\mathrm{I}\Delta_0+\tau+\Omega_n}(k, \phi(\bar{a}_1, \ldots, \bar{a}_l))$, then clearly $\phi(c_{a_1}, \ldots, c_{a_l}) \in T$. In particular we have (a) $\mathrm{I}\Delta_0 + \tau + \Omega_n \subseteq T$, (b) for all $a_1, \ldots, a_l \in M$, and all R_1^+ formulas $\phi(x_1, \ldots, x_l)$, if $M \vDash \phi(a_1, \ldots, a_l)$ then $\phi(c_{a_1}, \ldots, c_{a_l}) \in T$ (by 6.4, since $n \geqslant 1$), and (c) for all $a \in M$, $\exists z\ (z = B_{n+1}(\bar{a})) \in T$ (by 8.1). The result now follows upon taking $M' \vDash T$, setting $M^* = M' \restriction L$, and identifying $a \in M$ with $c_a^{M'} \in M^*$. □

We can now deduce a more general version of 3.3.:

8.3. Theorem. *If, in* 8.2, *τ is a Π_1 sentence, then M^* can be chosen to satisfy* exp.

Proof. This is clear from 8.2 since by (iii) there is an initial segment of the M^* given by that theorem which contains M (as a $\subseteq_+$-substructure) and which is closed under exponentiation. □

The following, which contains the $\Rightarrow$ direction of 3.4, is immediate from 8.3:

8.4. Theorem. *Suppose $\phi(\boldsymbol{x})$ is an R_1^+ formula, τ a Π_1 sentence and that* $\mathrm{I}\Delta_0 + \tau + \exp \vdash \forall \boldsymbol{x} \neg\phi(\boldsymbol{x})$. *Then for some $k \in \omega$,*

$$\mathrm{I}\Delta_0 + \Omega_1 + \mathrm{Con}(\mathrm{I}\Delta_0 + \tau, k) \vdash \forall \boldsymbol{x} \neg\phi(\boldsymbol{x}).$$

The converse of 8.4 obviously follows from the following

8.5. Proposition. *For all $k \in \omega$ and Π_1 sentences σ of L,*

$$\mathrm{I}\Delta_0 + \sigma + \exp \vdash \mathrm{Con}(\mathrm{I}\Delta_0 + \sigma, \mathbf{k}).$$

Proof. This proof uses the proposition of 7.15, but since we are not going to prove that result in this paper we sketch another proof of 8.5 below.

Suppose then that $k \in \omega$, τ is a Π_1 sentence of L and that M is a countable, non-standard model of $\mathrm{I}\Delta_0 + \sigma + \exp + \neg\mathrm{Con}(\mathrm{I}\Delta_0 + \sigma, k)$. We deduce a contradiction.

Let $p \in M$ be a k-proof of $\mathbf{0} = \mathbf{1}$ from $\mathrm{I}\Delta_0 + \sigma$, so that there is a finite set $\{\phi_1(x_1, \ldots, x_l, y), \ldots, \phi_r(x_1, \ldots, x_l, y)\}$ of formulas of L such that p is of the form $\phi_{i_0}(\tau_{0,1}, \ldots, \tau_{0,l}, y), \ldots, \phi_{i_\alpha}(\tau_{\alpha,1}, \ldots, \tau_{\alpha,l}, y)$ where $\alpha \in M$, $i_1, \ldots, i_l \in \{1, \ldots, r\}$, $\tau_{j,h}$ is a closed term (in M) for $j = 0, \ldots, \alpha$, $h = 1, \ldots, l$ and ϕ_{i_α} is $\mathbf{0} = \mathbf{1}$. Since the sequences of i_j's and $\tau_{j,h}$'s are coded in M we may consider the following formula, $\Psi(x)$, of L (with parameters from M):

$$\Psi(x) \overset{\text{def}}{\Leftrightarrow} \forall j \leqslant x \left(j \leqslant \alpha \rightarrow \bigvee_{m=1}^{r} (i_j = m \wedge \forall \boldsymbol{y}\ \phi_m(\mathrm{Val}(\tau_{j,1}), \ldots, \mathrm{Val}(\tau_{j,l}), \boldsymbol{y}))\right).$$

(cf. 4.14(7) for the definition of Val).

Now choose $n \in \omega$ such that Ψ is a Σ_n formula, and let $b \in M$ be large compared with all the parameters appearing in Ψ. Let k be as in the proposition of 7.15 and choose $a \in M$, $a \geqslant B_k(b)$. Let I be the initial segment of M given by this proposition, so that $b \in I$ (and hence p and all the parameters of Ψ are in I), $I \models \mathrm{I}\Delta_0 + \sigma$ and

$$I \models (\Psi(0) \wedge \forall x\, (\Psi(x) \rightarrow \Psi(x'))) \rightarrow \Psi(b).$$

Now it is easy to show, using the fact that p is a proof in M and hence in I, that

$$I \models \Psi(0) \wedge \forall x\, (\Psi(x) \rightarrow \Psi(x')),$$

and so $I \models \Psi(b)$. In particular, since $b > \alpha$,

$$I \models \forall y\, \phi_{i_\alpha}(\mathrm{Val}(\tau_{\alpha,1}), \ldots, \mathrm{Val}(\tau_{\alpha,l}), y),$$

i.e. $I \models \mathbf{0} = \mathbf{1}$, which is absurd. $\square$

We now sketch a proof of 8.5 that does not use (at least explicitly) the proposition of 7.15.

Suppose $k \in \omega$, $M \models \mathrm{I}\Delta_0 + \sigma + \exp$ (σ a Π_1 sentence of L) and that $p \in M$ is a k-proof of $\mathbf{0} = \mathbf{1}$ from $\mathrm{I}\Delta_0 + \sigma$. Then p only uses a finite set of axioms, say T, from $\mathrm{I}\Delta_0 + \sigma$, and these are all Π_1 sentences. Now we can find a cut-free proof, p^*, in M of $\mathbf{0} = \mathbf{1}$ from T because while $\mathrm{I}\Delta_0 + \exp$ cannot prove the full cut-elimination theorem (more comments on this later), it can prove such a theorem for k-proofs (with cut), where k is standard. (The p^* here will have length approximately $B_k(|p|)$.) The property we require of p^* is that all formulas occurring in it are subformulas of formulas in $T \cup \{\mathbf{0} = \mathbf{1}\}$, and hence Π_1. It is now an easy matter to find a suitably large $a \in M$, and a truth definition in M such that one can prove inductively that $\sigma^{\leqslant a}$ is true in M whenever σ occurs in p^*, where $\sigma^{\leqslant a}$ denotes the result of bounding all unbounded universal quantifiers in σ by a. This induction is possible in M because the $\sigma^{\leqslant a}$ are all Δ_0 formulas.

We summarize 8.4 and 8.5 in

8.6. Theorem. *Suppose $\phi(x)$ is an R_1^+ formula of L and that σ is any Π_1 sentence of L. Then* $\mathrm{I}\Delta_0 + \sigma + \exp \vdash \forall x \neg\phi(x)$ *if and only if*

$$\mathrm{I}\Delta_0 + \Omega_1 + \{\mathrm{Con}(\mathrm{I}\Delta_0 + \sigma, \mathbf{k}) : k \in \omega\} \vdash \forall x \neg\phi(x).$$

In particular, the two theories here have the same universal consequences.

As an application of 8.5 we characterize those Π_1 sentences σ such that $\mathrm{I}\Delta_0 + \sigma$ is interpretable in $\mathrm{I}\Delta_0$ by an initial formula. They are exactly those provable from $\mathrm{I}\Delta_0 + \exp$, which clearly follows from the following

8.7. Corollary. *Let $\psi(x)$ be Δ_0. Then there is an initial formula $\phi(x)$ such that $I\Delta_0 \vdash \forall x\, (\phi(x) \rightarrow \psi(x))$ if and only if* $\mathrm{I}\Delta_0 + \exp \vdash \forall x\, \psi(x)$.

Proof. Assume that $I\Delta_0 + \exp \vdash \forall x\, \psi(x)$. Then by 7.6

$$I\Delta_0 \vdash \forall x\, (\exists z\, z = B_{n+1}(x) \rightarrow \psi(x)) \quad \text{for some } n.$$

Hence for $Q_{n+1}(x)$ as in the proof of 8.1, $I\Delta_0 \vdash \forall x\, (Q_{n+1}(x) \rightarrow \psi(x))$.

Conversely suppose that $I\Delta_0 \vdash \forall x\, (\phi(x) \rightarrow \psi(x))$ for some initial formula $\phi(x)$. Let M be a countable model of $I\Delta_0 + \exp$ and suppose that $\forall x\, \psi(x)$ fails in M, say $M \models \neg\psi(a)$. Then, since exp holds in M it is easy to see that there is a k-proof of $\neg\psi(\bar{a})$ from $I\Delta_0$ in M for some $k \in \omega$. (See 2.7. Alternatively, by the Matijasevic–Robinson–Davis–Putnam theorem, which by results of [4] is provable in $I\Delta_0 + \exp$, there are polynomials $p(x, y)$, $q(x, y)$ over the natural numbers such that

$$I\Delta_0 + \exp \vdash \forall x\, (\neg\psi(x) \leftrightarrow \exists y\, (p(x, y) = q(x, y))).$$

Hence in M, $p(a, \boldsymbol{b}) = q(a, \boldsymbol{b})$ for some $\boldsymbol{b}$ and as in 6.4 there is a k-proof of $p(\bar{a}, \bar{\boldsymbol{b}}) = q(\bar{a}, \bar{\boldsymbol{b}})$ from $I\Delta_0$ in M.) But since $I\Delta_0 \vdash \phi(0) \wedge \forall x\, (\phi(x) \rightarrow \phi(x+1))$ we can show by induction on $b \in M$ that for some (fixed) $k_1 \in \omega$ there is a k_1-proof of $\phi(\bar{b})$ from $I\Delta_0$. In particular there is a k_1-proof of $\phi(\bar{a})$. Combining this with the standard proof of $\forall x\, (\phi(x) \rightarrow \psi(x))$ from $I\Delta_0$ gives a k_2-proof in M of $\psi(\bar{a})$ from $I\Delta_0$, for some $k_2 \in \omega$. Thus M contains a k_3-proof of $\psi(\bar{a}) \wedge \neg\psi(\bar{a})$ from $I\Delta_0$, for some $k_3 \in \omega$, and hence one of $\mathbf{0} = \mathbf{1}$. Thus $M \models \neg\mathrm{Con}(I\Delta_0, k_3)$ contradicting 8.5. □

Using the remarks after the proof of 7.5 one can now deduce a characterization of the Π_1 sentences interpretable in Robinson's Arithmetic Q. The precise result, the proof of which we leave to the reader, is as follows:

8.8. Corollary. *Let $\psi(x)$ be Δ_0. Then $I\Delta_0 + \exp \vdash \forall x\, \psi(x)$ if and only if there is a formula $\phi(x)$ such that $Q \vdash \phi(0) \wedge \forall x\, (\phi(x) \rightarrow \phi(x+1))$ and $Q \vdash \forall x\, (\phi(x) \rightarrow \psi(x))$.*

In view of 8.6 it is natural to ask whether the theories $I\Delta_0 + \Omega_1 + \mathrm{Con}(I\Delta_0, \mathbf{k})$ increase in strength as k increases. Indeed, none of our results so far rule out the possibility that $\mathrm{Con}(I\Delta_0, \mathbf{k})$ is provable from $I\Delta_0 + \Omega_1$ for all $k \in \omega$ and it is this question that we investigate now. We shall first prove 3.6 for which we require the (formalized) notion of a cut-free proof and the tableau method seems most convenient here.

8.9. Tableaux proofs. Let T be a set of sentences. We say that a sequence of sets of sets of formulas $\Gamma_0, \Gamma_1, \ldots, \Gamma_s$ is a tableau proof from T of a contradiction if the following conditions hold:

(i) For each $X \in \Gamma_s$, there is an atomic θ such that $\theta \in X$ and $\neg\theta \in X$.

(ii) $X \in \Gamma_0$ implies $X \subseteq T \cup \{$the logical equality axioms$\}$.

(iii) For each $X \in \Gamma_i$ with $i < s$ one of the following holds:

(α) $X \in \Gamma_{i+1}$,
(β) $X \cup \{\theta\} \in \Gamma_{i+1}$ for some $\neg\neg\theta \in X$,
(γ) $X \cup \{\neg\theta_1\}, X \cup \{\theta_2\} \in \Gamma_{i+1}$ for some $(\theta_1 \rightarrow \theta_2) \in X$,
(δ) $X \cup \{\theta_1, \neg\theta_2\} \in \Gamma_{i+1}$ for some $\neg(\theta_1 \rightarrow \theta_2) \in X$,
(ε) $X \cup \{\theta(t)\} \in \Gamma_{i+1}$ for some $\forall x\, \theta(x) \in X$ and some term t which is freely substitutable for x in $\theta(x)$,
(ζ) $X \cup \{\neg\theta(y)\}$ for some $\neg\forall x\, \theta(x) \in X$ and some variable y which does not occur in any formula in X.

(iv) For each $Y \in \Gamma_{i+1}$ with $i < s$ there is an $X \in \Gamma_i$ such that Y is obtained from X by one of the rules (iii)(α)–(ζ).

It is well known that T is inconsistent in the usual sense just if there is a tableau proof from T of a contradiction (notice that the usual logical equality axioms have been implicitly included in T). The advantage of such tableau proofs is that they only contain subformulas of the sentences in T. The disadvantage however is that they are in general 'iteratedly exponentially longer' than conventional proofs.

We denote by L^* the language of arithmetic where successor, addition and multiplication are (2-, 3- and 3-place) relation symbols. Thus, the only terms of L^* are variables and $\mathbf{0}$. We denote by $\mathrm{I}\Delta_0^*$ the obvious reformulation of $\mathrm{I}\Delta_0$ in L^*, where we include in $\mathrm{I}\Delta_0^*$ the sentences expressing that the successor, addition and multiplication relations are total functions. Notice that these latter three axioms are Π_2 whereas all the other axioms of $\mathrm{I}\Delta_0^*$ remain Π_1. In general, if ϕ *is a formula of* L we denote its natural reformulation in L^* by ϕ^* and we may suppose that this translation is carried out so as to preserve (unbounded) quantifier complexity. Now using the methods of Sections 4 and 5 it is routine to express all formalized syntactic notions of the language L^* using R_1^+ formulas (of L). In particular for suitable theories T of L^* (e.g. for any finite extension of $\mathrm{I}\Delta_0^*$) there is an R_1^+ formula (of L), $\mathrm{Tabinconpr}(T, x)$ adequately expressing, in any model of $\mathrm{I}\Delta_0 + \Omega_1$, that "$x$ is a tableau proof from T of a contradiction". We denote by $\mathrm{Tabincon}(T)$ the sentence $\exists x\, \mathrm{Tabiconpr}(T, x)$.

8.10. Lemma. *Let $r \in \omega$ and σ be any Σ_2 sentence of L. Then*

$$\mathrm{I}\Delta_0 + \sigma + \exp \vdash \neg\mathrm{Tabincon}(\mathrm{I}\Delta_0^* + \sigma^* + \Omega_r^*).$$

Proof. Let $M \models \mathrm{I}\Delta_0 + \sigma + \exp$ and suppose that $M \models \mathrm{Tabincon}(\mathrm{I}\Delta_0^* + \sigma^* + \Omega_r^*)$. The idea of the proof is similar to that of the second proof of 8.5 in that we shall use the truth definition for Δ_0 formulas given in [7]. Let us denote by M^* the structure M regarded as an L^*-structure, and by a^* (for $a \in M$) the L^*-structure of M^* with domain $\{\alpha \in M^* : M^* \models \alpha < a\}$. Then (since M satisfies exp) the results of [7] imply that for any b, $c \in M$, there is a Δ_0 formula $T_{b,c}(x, y)$ (with parameters from M) asserting that "if $x = \ulcorner\theta(v_1, \ldots, v_t)\urcorner$ is any L^* formula with $x < c$ and $y = \langle b_1, \ldots, b_t\rangle$ is a sequence of elements of b^*, then $\theta(b_1, \ldots, b_t)$ is true in b^*." Of course this only really makes sense if $x = \ulcorner\theta(v_1, \ldots, v_t)\urcorner$ is standard, but the point is that $T_{b,c}(x, y)$ satisfies the usual truth conditions even for non-standard $x < c$. In particular it follows that any Π_1 axiom of $\mathrm{I}\Delta_0^*$ with Gödel number less than c will be 'true' in b^*. We shall write $b^* \vDash \theta(b_1, \ldots, b_t)$ for $T_{b,c}(x, y)$ whenever $t, x, y \in M$ satisfy $x = \ulcorner\theta(v_1, \ldots, v_t)\urcorner$, $x < c$, $y = \langle b_1, \ldots, b_t\rangle$, $b_i < b$ for $i = 1, \ldots, t$.

To return to the matter at hand let $a \in M$ satisfy $M \vDash \forall z\, \lambda(a, z)$ where σ is $\exists x\, \forall z\, \lambda(x, z)$ and suppose $\Gamma_1, \ldots, \Gamma_s$ is a tableau proof (in M) from $\mathrm{I}\Delta_0^* + \sigma^* + \Omega_r^*$ of a contradiction. Let $c \in M$ be larger than any formula occurring in any set in any Γ_i and set $b = \omega_r^{(s+1)}(a + 2)$. (It is routine to check that the function $(x, y) \to \omega_r^{(x)}(y) = \omega_r$ applied x times to y, is well defined and total in M.) Now for each $i < s$ and $X \in \Gamma_i$ we define in M a function $F_{i,X}$ with domain the set of variables occurring in formulas in X and range bounded by b, by induction on i as follows. If $i = 0$ then $F_{i,X}$ is empty. For u a variable in (some formula in) $Y \in \Gamma_{i+1}$ pick (by 8.9(iv)) $X \in \Gamma_i$ such that Y is derived from X by one of 8.9(iii)(α)–(ζ). If u appears in X set $F_{i+1,Y}(u) = F_{i,X}(u)$. If u occurs by (ζ), say $Y = X \cup \{\neg\theta(u, x_1, \ldots, x_p)\}$ where $\neg\forall x\, \theta(x, x_1, \ldots, x_p) \in X$, set

$$F_{i+1,Y}(u) = \begin{cases} \text{the least } d < b \text{ such that } b^* \vDash \neg\theta(d, F_{i,X}(x_1), \ldots F_{i,X}(x_p)), \\ 0 \quad \text{if no such } d \text{ exists.} \end{cases}$$

In all other cases set $F_{i+1,Y}(u) = 0$.

It is easy to check that the above definition can be carried out in M and (by induction on i) that the following hold for each $i \leq s$:

(1) $\forall X \in \Gamma_i$, $\mathrm{Range}(F_{i,X}) \subseteq \{\alpha \in M : M \vDash \alpha < \omega_r^{(i+1)}(a + 2)\}$,

(2) $\exists X \in \Gamma_i$ such that for all formulas $\theta(x_1, \ldots, x_p)$ in X which are either Σ_1 or Π_1, $b^* \vDash \theta(F_{i,X}(x_1), \ldots, F_{i,X}(x_p))$.

(For (1), note that the only time 8.9(iii)(ζ) can be used to eliminate an unbounded quantifier is on a formula of the form $\neg\forall x\, \neg f(y) = x$, where $f(y)$ is y' or $\omega_r(y)$, or $\neg\forall x\, \neg f(y_1, y_2) = x$, where $f(y_1, y_2)$ is $y_1 + y_2$ or $y_1 \cdot y_2$. Thus the value given to a new variable is at most ω_r of the value given to some old variable. Further, since the only terms of L^* are variables or $\mathbf{0}$, the values given to new variables introduced by 8.9(iii)(ε) or by application of 8.9(iii)(ζ) to a bounded quantifier will always be greater than values given to old variables (or will be 0). (2) follows from the properties of our truth definition.)

But now (2) clearly contradicts 8.9(i) which establishes the lemma. $\square$

We can now establish 3.6:

8.11. Theorem. *Let $r \in \omega$ and suppose that σ is a Σ_2 sentence and $\mathrm{I}\Delta_0 + \sigma + \exp$ is consistent. Then there is an R_1^+ formula $\phi(x)$ such that*

$$\mathrm{I}\Delta_0 + \sigma + \exp \vdash \forall x\, \neg\phi(x) \quad \textit{but} \quad \mathrm{I}\Delta_0 + \sigma + \Omega_r \nvdash \forall x\, \neg\phi(x).$$

Proof. By the fixed point theorem (see, e.g. [1]) there is a sentence ψ of L such that $I\Delta_0 \vdash \psi \leftrightarrow \text{Tabincon}(I\Delta_0^* + \sigma^* + \Omega_r^* + \psi^*)$. Further ψ has the form $\exists x\, \phi(x)$, with $\phi(x)$ R_1^+ (because the formula $\text{Tabincon}(I\Delta_0^* + \sigma^* + \Omega_r^* + y^*)$ does). We claim $I\Delta_0 + \sigma + \exp \vdash \neg\psi$ but that $I\Delta_0 + \sigma + \Omega_r \nvdash \neg\psi$, so completing the proof. To see that $I\Delta_0 + \sigma + \exp \vdash \neg\psi$ suppose $M \vDash I\Delta_0 + \exp + \psi$. Then by 8.10 (after combining σ, ψ into a single Σ_2 sentence) we have $M \vDash \neg\text{Tabincon}(I\Delta_0^* + \sigma^* + \Omega_r^* + \psi^*)$ and hence $M \vDash \neg\psi$, a contradiction.

Finally, suppose $I\Delta_0 + \sigma + \Omega_r \vdash \neg\psi$. Then $I\Delta_0 + \sigma + \Omega_r + \psi$ would be inconsistent and hence, clearly, $I\Delta_0^* + \sigma^* + \Omega_r^* + \psi^*$ would be also. But then there would be a tableau proof from $I\Delta_0^* + \sigma^* + \Omega_r^* + \psi^*$ of a contradiction, and this would be just a standard proof and so it would exist in any model of $I\Delta_0$. Hence $I\Delta_0 \vdash \text{Tabincon}(I\Delta_0^* + \sigma^* + \Omega_r^* + \psi^*)$, and so $I\Delta_0 \vdash \psi$ and therefore $I\Delta_0 + \sigma + \Omega_r$ would be inconsistent, which clearly contradicts the assumption that $I\Delta_0 + \sigma + \exp$ is consistent. □

8.12. Remark. Notice that in 8.11 the formula ϕ depends on r. Indeed, it is an open problem as to whether an R_1^+ formula, $\phi(x)$, exists with the property that $I\Delta_0 + \exp \vdash \forall x\, \neg\phi(x)$ but $I\Delta_0 + \{\Omega_r : r \in \omega\} \nvdash \forall x\, \neg\phi(x)$. The difficulty in generalizing 8.10 is that we know of no function f with a Δ_0 graph which grows reasonably smoothly and is such that every model of $I\Delta_0 + \exp$ is closed under the iterated function $(x, y) \to f^{(x)}(y)$, except those that are majorized by ω_r (for some fixed r). (Notice that this property of the ω_r was used crucially in the proof of 8.10.)

8.13. Corollary. *Let k, $r \in \omega$. Then there is $n \in \omega$ such that*

$$I\Delta_0 + \Omega_r + \text{Con}(I\Delta_0, \mathbf{k}) \nvdash \text{Con}(I\Delta_0, \mathbf{n}).$$

Proof. Since $\text{Con}(I\Delta_0, \mathbf{k})$ is a Π_1 sentence (or, rather, may easily be put in Π_1 form) there is an R_1^+ formula $\phi(x)$ such that $I\Delta_0 + \exp + \text{Con}(I\Delta_0, \mathbf{k}) \vdash \forall x\, \neg\phi(x)$ but $I\Delta_0 + \Omega_r + \text{Con}(I\Delta_0, \mathbf{k}) \nvdash \forall x\, \neg\phi(x)$ (by 8.11 with $\sigma = \text{Con}(I\Delta_0, \mathbf{k})$). Since by 8.5 $I\Delta_0 + \exp \vdash \text{Con}(I\Delta_0, \mathbf{k})$, we have that $I\Delta_0 + \exp \vdash \forall x\, \neg\phi(x)$. Therefore by 8.4, $I\Delta_0 + \Omega_1 + \text{Con}(I\Delta_0, \mathbf{n}) \vdash \forall x\, \neg\phi(x)$ for some $n \in \omega$. It follows that, since we may assume $r > 0$, $I\Delta_0 + \Omega_r + \text{Con}(I\Delta_0, \mathbf{k}) \nvdash \text{Con}(I\Delta_0, \mathbf{n})$. □

We can now prove the first part of Theorem 3.5.:

8.14. Corollary. $I\Delta_0 + \exp \nvdash \text{Con}(I\Delta_0)$. *More generally, if σ is a Π_1 sentence and* $I\Delta_0 + \exp + \sigma$ *is consistent, then* $I\Delta_0 + \sigma + \exp \nvdash \text{Con}(I\Delta_0 + \sigma)$. *Further,* $I\Delta_0 + \exp \nvdash \text{Con}(Q)$, *where Q is Robinson's arithmetic.*

Proof. Assume $I\Delta_0 + \exp \vdash \text{Con}(I\Delta_0)$. Then since $\text{Con}(I\Delta_0)$ has the form $\forall x\, \neg\phi(x)$ with $\phi(x)$ R_1^+, we have, by 8.4, that

$$I\Delta_0 + \Omega_1 + \text{Con}(I\Delta_0, \mathbf{k}) \vdash \text{Con}(I\Delta_0), \quad \text{for some } k \in \omega.$$

But then clearly

$$\mathrm{I}\Delta_0 + \Omega_1 + \mathrm{Con}(\mathrm{I}\Delta_0, \mathbf{k}) \vdash \mathrm{Con}(\mathrm{I}\Delta_0, \mathbf{n}) \quad \text{for all } n \in \omega$$

which contradicts 8.13. The generalization follows in the same way by generalizing 8.13 according to 8.11 and using the full strength of 8.4. The result concerning Con(Q) follows from the remarks after the proof of 7.5. □

Recall from 1.1 that one of our aims in this paper was to investigate the Π_1 consequences of $\mathrm{I}\Delta_0 + \exp$ modulo $\mathrm{I}\Delta_0$ (or modulo $\mathrm{I}\Delta_0 + \Omega_r$). Let U_1 denote the set of sentences of the form $\forall x \neg\phi(x)$ for $\phi(x)$ R_1^+. Then certainly every sentence in U_1 is equivalent, modulo $\mathrm{I}\Delta_0 + \Omega_1$, to one in Π_1. Further, every sentence in Π_1 is equivalent, modulo $\mathrm{I}\Delta_0 + \exp$, to one in U_1 since the Matijasevic–Robinson–Davis–Putman theorem is provable in $\mathrm{I}\Delta_0 + \exp$ (see [4]) and U_1 certainly includes all purely universal sentences. Unfortunately, since we do not know whether this theorem is provable in $\mathrm{I}\Delta_0 + \Omega_r$ (for any $r \in \omega$) (cf. 1.3), we do not know whether U_1 and Π_1 are equivalent modulo $\mathrm{I}\Delta_0 + \Omega_r$ and it is this difficulty that prevents us from characterizing the Π_1 consequences of $\mathrm{I}\Delta_0 + \exp$ (modulo $\mathrm{I}\Delta_0 + \Omega_r$). However, we have now completed our characterization of the U_1 consequences of $\mathrm{I}\Delta_0 + \exp$ (modulo $\mathrm{I}\Delta_0 + \Omega_r$) which we now summarize:

8.15. Theorem. *The* U_1 *sentences* $\mathrm{Con}(\mathrm{I}\Delta_0, \mathbf{k})$ $(k \in \omega)$ *exhaust the* U_1 *consequences of* $\mathrm{I}\Delta_0 + \exp$ *in the presence of the theory* $\mathrm{I}\Delta_0 + \Omega_1$. *Further, for each* $r \in \omega$ *there is an increasing sequence* $k_0 < k_1 < \cdots$ *of natural numbers such that* $s < t$ *implies* $\mathrm{I}\Delta_0 + \Omega_r \nvdash \mathrm{Con}(\mathrm{I}\Delta_0, \mathbf{k}_s) \rightarrow \mathrm{Con}(\mathrm{I}\Delta_0, \mathbf{k}_t)$. *In particular* $\mathrm{I}\Delta_0 + \exp \nvdash \mathrm{Con}(\mathrm{I}\Delta_0)$.

8.16. An application. Let us write $E_n(x)$ for $e_n(x, \ldots, x)$ (cf. 2.4). Now it is well known that for every proof, p, in a (natural) formal system with a cut rule there is a proof, p_1, of the same sentence in any natural cut-free system. However, as we have already mentioned, it is also known that there is no fixed $n \in \omega$ such that $\ulcorner p_1 \urcorner$ can always be chosen less than $E_n(\ulcorner p \urcorner)$. It is perhaps not surprising that this latter result follows easily from 8.14 in view of the techniques used in our proof of that theorem. To see this (at least for the tableau system of 8.9) let $n \in \omega$ and define the sentence σ (of L) by

$$\sigma_n = \forall x, y, z\ ((P(x, y) \wedge z = E_n(y)) \rightarrow \exists u \leq z\ \mathrm{Tabinconpr}(\mathrm{Neg}(x), u)),$$

where $P(x, y)$ is an R_1^+ formula naturally expressing "y is a proof of the L^*-sentence x (from non-logical axioms) in $\mathfrak{F}$", where $\mathfrak{F}$ is some formal system with language L^* and with a cut rule (say the obvious modification of the system of Section 5, where the cut rule is, of course, modus ponens), and where, if $x = \ulcorner \phi \urcorner$, then $\mathrm{Neg}(x) = \ulcorner \neg\phi \urcorner$. Notice that σ_n is a Π_1 sentence.

We must show that σ_n is false (in the standard model). Indeed, we shall show that $\mathrm{I}\Delta_0 + \sigma_n + \exp$ is inconsistent. For suppose $M \models \mathrm{I}\Delta_0 + \sigma_n + \exp$. Let τ be a sentence of L^* in M that is a conjunction of σ_n^* and (M-) finitely many axioms of $\mathrm{I}\Delta_0^*$ (cf. 8.9), and let $\tau_1 = (\tau \rightarrow (\mathbf{0}=\mathbf{1})^*)$. Then we claim that $M \models \forall y \neg P(\tau_1, y)$. For if $p \in M$ and $M \models P(\tau_1, p)$ then, since $M \models \sigma_n \wedge \exists z\ (z = E_n(p))$ we would have $M \models \mathrm{Tabincon}(\mathrm{Neg}(\tau_1))$ and hence $M \models \mathrm{Tabincon}(\mathrm{I}\Delta_0^* + \sigma_n^* + \neg(\mathbf{0}=\mathbf{1})^*)$, which contradicts 8.10. It now follows that $M \models \mathrm{Con}(\mathrm{I}\Delta_0 + \sigma_n)$, since any proof in M of $\mathbf{0}=\mathbf{1}$ from $\mathrm{I}\Delta_0 + \sigma_n$ in the system we introduced in Section 5 (in the language L) could easily be converted to one of $(\mathbf{0}=\mathbf{1})^*$ from a (M-) finite subset of $\mathrm{I}\Delta_0^* + \sigma_n^*$ (in the system $\mathfrak{F}$). But we have now shown that $\mathrm{I}\Delta_0 + \sigma_n + \exp \vdash \mathrm{Con}(\mathrm{I}\Delta_0 + \sigma_n)$ and so, by 8.14, $\mathrm{I}\Delta_0 + \sigma_n + \exp$ is inconsistent as required.

8.17. We shall now use a version of the cut elimination theorem combined with a model-theoretic argument along the lines of 8.1 to establish the last part of Theorem 3.5.

We first define the super-exponential function $\mathrm{supexp}(x, y)$ by:

$$\mathrm{supexp}(x, 0) = x; \qquad \mathrm{supexp}(x, y+1) = x^{\mathrm{supexp}(x,y)}.$$

It is easy to show that the graph of supexp can be expressed by a Δ_0 formula and that the recursive defining equations can be proved (in $\mathrm{I}\Delta_0$) to hold whenever the function is defined. Further, if $M \models \mathrm{I}\Delta_0$, then for all $a \in M \setminus \{0, 1\}$ the set $\{b \in M : M \models \exists z\ z = \mathrm{supexp}(a, b)\}$ is an initial segment of M which will be closed under successor if and only if $M \models \exp$.

Now inspection of the proof of the cut elimination theorem reveals that it can be carried out in $\mathrm{I}\Delta_0$ provided one has certain superexponentials to work with. More precisely we have the following result.

8.18. Lemma. *Suppose $M \models \mathrm{I}\Delta_0 + \Omega_1$ and that $p \in M$ is a proof in the system introduced in Section 5. Supppose $b \in M$ and $M \models \exists z\ (z = \mathrm{supexp}(B, p) \wedge z \leqslant b)$. Then there is a tableau proof, p_1, from $\neg\tau^*$ of a contradiction in M such that $p_1 \leqslant b$, where τ is the sentence $(\tau_1 \rightarrow \tau_2)$ and τ_1, τ_2 are respectively the conjunction of the nonlogical axioms and conclusion of p. In particular if p is a proof of $\mathbf{0}=\mathbf{1}$ from* $\mathrm{I}\Delta_0$, *then (by* 8.10) $M \models \neg\exp$.

8.19. Theorem. $\mathrm{I}\Delta_0 + \exp + \mathrm{Con}(\mathrm{I}\Delta_0) \nvdash \mathrm{Con}(\mathrm{I}\Delta_0 + \exp)$.

Proof. We first denote by $K_0(x)$ the formula $\exists z\ z = \mathrm{supexp}(\mathbf{B}, x)$. Then, as remarked in 8.17, if $M \models \mathrm{I}\Delta_0 + \exp$ we have that $K_0^M \subseteq_e M$ (i.e., K_0^M is an initial segment of M) and K_0^M is closed under successor.

We now define formulas $K_i(x)$, $L_i(x)$, $M_i(x)$ and $N_i(x)$ (for $i \in \omega$), which will satisfy $K_i^M \supseteq_e L_i^M \supseteq_e M_i^M \supseteq_e N_i^M \supseteq_e K_{i+1}^M$ for all $i \in \omega$ and $M \models \mathrm{I}\Delta_0 + \exp$, by induction on i as follows:

$$L_i(x) \overset{\text{def}}{\Leftrightarrow} \forall y\,(K_i(y) \to K_i(x+y)),$$

$$M_i(x) \overset{\text{def}}{\Leftrightarrow} \forall y\,(L_i(y) \to L_i(x \cdot y)),$$

$$N_i(x) \overset{\text{def}}{\Leftrightarrow} \forall y\,(M_i(y) \to M_i(y^{|x|})),$$

$$K_{i+1}(x) \overset{\text{def}}{\Leftrightarrow} \exists y\,(N_i(y) \wedge y = \mathbf{B}^x).$$

As in previous arguments it is straightforward to check that these formulas have the above initial segment property, that, for any $i \in \omega$ and $M \models \mathrm{I}\Delta_0 +$ exp, N_i^M is closed under ω_1^M (i.e., $N_i^M \models \Omega_1$) and that

(α) $\forall a \in N_i^M\ N_0^M \models \exists z\ z = B_i(a)$ (cf. the statement of 8.1).

Further, just as in the proof of 8.1, one can show

(β) $\forall a \in M$, there is a proof in M of the sentence $N_i(\bar{a})$ from $\mathrm{I}\Delta_0 + \exp$.

Now suppose for contradiction, that $\mathrm{I}\Delta_0 + \exp + \mathrm{Con}(\mathrm{I}\Delta_0) \vdash \mathrm{Con}(\mathrm{I}\Delta_0 + \exp)$. Then by 7.6 there is $m \in \omega$ such that

(γ) $\mathrm{I}\Delta_0 + \mathrm{Con}(\mathrm{I}\Delta_0) \vdash \forall x\,(\exists z\ z = B_{m+1}(x) \to \neg\mathrm{Proof}_{\mathrm{I}\Delta_0+\exp}(\mathbf{k}, x)$ where $k = \ulcorner \mathbf{0} = \mathbf{1} \urcorner$.

We now claim that

(δ) If $M \models \mathrm{I}\Delta_0 + \exp$, then $N_{m+1}^M \models \mathrm{I}\Delta_0 + \Omega_1 + \mathrm{Con}(\mathrm{I}\Delta_0 + \exp)$.

For we have already observed that $N_{m+1}^M \models \mathrm{I}\Delta_0 + \Omega_1$ so it only remains to show that $N_{m+1}^M \models \mathrm{Con}(\mathrm{I}\Delta_0 + \exp)$ for which it suffices to show, by (α) and (γ) (and the fact that the formula $\mathrm{Proof}_{\mathrm{I}\Delta_0+\exp}$ is preserved in initial segments of models of $\mathrm{I}\Delta_0 + \Omega_1$), that $N_0^M \models \mathrm{Con}(\mathrm{I}\Delta_0)$. However since $p \in N_0^M \subseteq_e K_0^M$ implies $M \models \exists z\ z =$ supexp(B, p), this follows immediately from 8.18, and (δ) is established.

Our idea now is to construct $\subseteq_+$-extensions of models of $\mathrm{I}\Delta_0 + \Omega_1 +$ $\mathrm{Con}(\mathrm{I}\Delta_0 + \exp)$ roughly along the same lines as in 8.3. However, we start the process in a different way. We first set out to define finite theories $S_0 \subseteq S_1 \subseteq \cdots$ as follows.

Let θ_i, $i \in \omega$, be a recursive enumeration of all sentences of L (so that there is a uniform definition of the function $i \mapsto \ulcorner \theta_i \urcorner$, that works at least for $i \in \omega$, in all models of $\mathrm{I}\Delta_0$) and define

$$S_0 = \{\exp\},$$

$$S_{i+1} = \begin{cases} S_i \cup \{\theta_i\} & \text{if } \mathrm{Con}(\mathrm{I}\Delta_0 + S_i + \theta_i), \\ S_i \cup \{\neg\theta_i\} & \text{otherwise.} \end{cases}$$

Clearly the (partial) function $x \mapsto$ (a suitable code for) S_x is definable by a formula of L, and will be total on (at least) ω ($\subseteq_e M$) in all $M \models \mathrm{I}\Delta_0 + \Omega_1$. Also, notice that while S_i^M (for $i \in \omega$, $M \models \mathrm{I}\Delta_0 + \Omega_1$) is a (code for a) finite set of standard L-sentences, this set will in general vary with M because different consistency statements may hold in different models M. However, we can assert that if M, $J \models \mathrm{I}\Delta_0 + \Omega_1$, $i \in \omega$, $M \subseteq_+ J$ and $S_i^M = S_i^J$, then $\theta_i \in S_{i+1}^J$ implies $\theta_i \in S_{i+1}^M$. For if $p \in M$ satisfies $M \models \mathrm{Proof}_{\mathrm{I}\Delta_0 + \theta_i + S_i^M}(\ulcorner \mathbf{0} = \mathbf{1} \urcorner, p)$ then, since this formula is R_1^+ and $M \subseteq_+ J$, we have $J \models \mathrm{Proof}_{\mathrm{I}\Delta_0 + \theta_i + S_i^M}(\ulcorner \mathbf{0} = \mathbf{1} \urcorner, p)$ and hence (since $S_i^M = S_i^J$) $J \models \neg\mathrm{Con}(\mathrm{I}\Delta_0 + S_i + \theta_i)$.

Notice also that if $M \models \mathrm{I}\Delta_0 + \Omega_1 + \mathrm{Con}(\mathrm{I}\Delta_0 + \exp)$, then $M \models \mathrm{Con}(\mathrm{I}\Delta_0 + S_i)$ for all $i \in \omega$.

Now by a suitable application of the fixed point theorem (see [1]), there is $k \in \omega$ such that

(ε) $\mathrm{I}\Delta_0 + \Omega_1 \vdash \theta_k \leftrightarrow \neg(\ulcorner \theta_k \urcorner \in S_{k+1})^{N_{m+1}}$. (The m here is as in (γ), (δ).)

(Informally, θ_k holds in a model $M \models \mathrm{I}\Delta_0 + \Omega_1$ just if carrying out the construction of the S_i's inside N_{m+1}^M would lead to θ_k not being in the set S_{k+1}.)

We now associate to each model M of $\mathrm{I}\Delta_0 + \Omega_1$ a k-tuple, s_M, of 0's and 1's as follows:

$$\text{For } i = 1, \ldots, k \quad s_M(i) = \begin{cases} 0 & \text{if } \ulcorner \theta_i \urcorner \in S_{i+1}^M, \\ 1 & \text{if } \ulcorner \neg\theta_i \urcorner \in S_{i+1}^M. \end{cases}$$

Our assertion above clearly implies

(ζ) For M, $J \models \mathrm{I}\Delta_0 + \Omega_1$, $M \subseteq_+ J$ implies $s_M \leqslant s_J$ (in lexicographic order).

Let M be an arbitrary countable model of $\mathrm{I}\Delta_0 + \Omega_1 + \mathrm{Con}(\mathrm{I}\Delta_0 + \exp)$. Our aim is to construct a countable model J of this same theory such that $M \subseteq_+ J$ and $s_M \neq s_J$. This gives the required contradiction, for repeating this construction $2^k + 1$ times would result in, by (ζ), a strictly increasing (in lexicographic order) sequence, of length $2^k + 1$, of k-tuples of 0's and 1's!

To this end we use the fact that $M \models \mathrm{Con}(\mathrm{I}\Delta_0 + S_{k+1}^M)$ and employ the same method as in the proof of 8.2 (except that now full consistency in M is used, rather than 'k-consistency $\forall k \in \omega$') to obtain a countable model $M^\dagger \models \mathrm{I}\Delta_0 + S_{k+1}^M$ (so $M^\dagger \models \exp$, since $\exp \in S_{k+1}^M$) such that $M \subseteq_+ M^\dagger$. Further, by (β) and properties of this construction, it follows that $M^\dagger \models N_{m+1}(a)$ for all $a \in M$. Let $J = N_{m+1}^{M^\dagger}$. Then $M \subseteq_+ J$ (because $M \subseteq J \subseteq_e M^\dagger$ and $M \subseteq_+ M^\dagger$) and $J \models \mathrm{I}\Delta_0 + \Omega_1 + \mathrm{Con}(\mathrm{I}\Delta_0 + \exp)$ (by (δ) applied to $M^\dagger$). Finally, (ε) clearly implies that $S_{k+1}^M \neq S_{k+1}^J$, so $s_M \neq s_J$ as required. □

We remark that the method of proof used here could also have been used to prove the first statement in 8.14 directly, i.e., without going via k-consistency statements, and we leave the reader to fill in the details of such an argument.

We conclude this paper with statements of results concerning iterated exponentiation which can be proved using easy modifications of the techniques we have been employing. Firstly, although we have seen that $\mathrm{I}\Delta_0 + \exp$ is not strong enough to prove Con($I\Delta_0$), it follows from 8.18 that $I\Delta_0 + \forall x, y\, \exists z\, z = \mathrm{supexp}(x, y)$ is sufficient. Indeed, rather more is true. Let Λ^f be the sentence

$$\forall x\, (\exists y\, y = \mathrm{supexp}(2, x) \rightarrow \exists y, z\, (z = f(x) \wedge y = \mathrm{supexp}(2, z)),$$

where f is a function with a Δ_0 graph. (Thus, for example, $\Lambda^{\text{successor}}$ is equivalent in $I\Delta_0$ to exp.) Then we have

8.20. Theorem. $I\Delta_0 + \forall x, y\, \exists z\, z = \mathrm{supexp}(x, y) \vdash \mathrm{Con}(I\Delta_0 + \Lambda^{e_1(2,\,-)})$.

The following theorem parallels 7.14 and 8.14 and is proved by similar methods.

8.21. (i) $I\Delta_0 + \Omega_1 + \mathrm{Con}(I\Delta_0 + \exp) \vdash \mathrm{Con}(I\Delta_0 + \Lambda^{\omega_n})$, *for each* $n \in \omega$.
(ii) $I\Delta_0 + \exp + \mathrm{Con}(I\Delta_0 + \exp) \nvdash \mathrm{Con}(I\Delta_0 + \Lambda^{e_1(2,\,-)})$.

References

[1] G. Boolos and R. Jeffrey, Computability and Logic, 2nd. edition (Cambridge Univ. Press, Cambridge, 1980).
[2] C. Dimitracopoulos, Matijasevič's theorem and fragments of arithmetic, Ph.D. Thesis, Manchester University, 1980.
[3] C. Dimitracopoulos and J.B. Paris, A note on the nondefinability of cuts, J. Symbolic Logic 48 (1983) 564–569.
[4] H. Gaifman and C. Dimitracopoulos, Fragments of Peano's arithmetic and the MRDP theorem, in: Logic and Algorithmic, Monographie No. 30 de L'Enseignement Mathématique (Genève, 1982) 187–206.
[5] E. Mendelson, Introduction to Mathematical Logic, 2nd edition, (Van Nostrand, New York) 494–508.
[6] R. Parikh, Existence and feasibility in arithmetic, J Symbolic Logic 36 (1971) 494–508.
[7] J.B. Paris and C. Dimitracopoulos, Truth definition for Δ_0 formulae, in: Logic and Algorithmic, Monographie No. 30 de L'Enseignement Mathématique (Genève, 1982) 319–329.
[8] J.B. Paris and A.J. Wilkie, Models of arithmetic and the rudimentary sets, Bull. Soc. Math. Belg. 33 (1) (1981) 157–169.
[9] J.B. Paris and A.J. Wilkie, Δ_0 sets and induction, in: Open Days in Model Theory and Set Theory, Proc. of a Conference held in September 1981 at Jadwisin, Poland (Leeds University, 1983) 237–248.
[10] A.J. Wilkie, Some results and problems on weak systems of arithmetic, in: A. MacIntyre et al., eds., Logic Colloquium '77 (North-Holland, Amsterdam, 1980) 285–296.
[11] A.J. Wilkie, Applications of complexity theory to Σ_0-definability problems in arithmetic, in: Model Theory of Algebra and Arithmetic, Lecture Notes in Math. 834 (Springer, Berlin, 1980) 363–369.
[12] A.R. Woods, Some problems in logic and number theory, and their connections, Ph.D. Thesis, Manchester University, 1981.